

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 08:53:35 ; Search time 4682 Seconds
(without alignments)
8160.955 Million cell updates/sec

Title: US-10-030-613-3
Perfect score: 934
Sequence: 1 agaatcattcttcgga.....gtataaccattatttagaa 934

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_wi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	934	100.0	934	6	AX074351	AX074351 Sequence
2	934	100.0	3159	6	AX420425	AX420425 Sequence
3	934	100.0	3261	9	AK090990	AK090990 Homo sapi
4	930.8	99.7	3240	9	AY034085	AY034085 Homo sapi
5	930.8	99.7	3253	9	AY034086	AY034086 Homo sapi
6	720.8	77.2	2523	9	BC038509	BC038509 Homo sapi
7	720.8	77.2	3184	6	AX329596	AX329596 Sequence
8	720.8	77.2	3184	6	AX333015	AX333015 Sequence
9	720.8	77.2	3184	6	AX365337	AX365337 Sequence
10	720.8	77.2	3184	6	AX420436	AX420436 Sequence
11	720.8	77.2	3184	9	HUMZAK14	D83407 ZAK1-4 mRNA
12	685	73.3	3295	10	AB061525	AB061525 Mus muscu
13	666.2	71.3	953	10	AF459023	AF459023 Rattus no
14	514.6	55.1	3080	10	BC047153	BC047153 Mus muscu
15	514.6	55.1	3263	10	BC049096	BC049096 Mus muscu
16	496.2	53.1	833	10	AF459022	AF459022 Rattus no
17	457.6	49.0	3227	10	AB061524	AB061524 Mus muscu
18	422	45.2	1021	6	AX420427	AX420427 Sequence
19	420	45.0	659	10	AF237887	AF237887 Mus muscu
20	413.2	44.2	594	6	AX365318	AX365318 Sequence
21	413.2	44.2	594	10	AF237791	AF237791 Mus muscu
22	374	40.0	128461	9	AL390741	AL390741 Human DNA
23	363	38.9	134443	2	AC015547	AC015547 Homo sapi
24	280.8	30.1	2827	9	BC035854	BC035854 Homo sapi
25	279.2	29.9	828	6	AX365330	AX365330 Sequence
26	279.2	29.9	828	6	AX420437	AX420437 Sequence
27	279.2	29.9	828	9	AF176116	AF176116 Homo sapi
28	268	28.7	776	10	AF237888	AF237888 Mus muscu
29	257.6	27.6	615	6	AX420432	AX420432 Sequence
30	255	27.3	226218	2	AC112622	AC112622 Rattus no
31	255	27.3	246030	2	AC126591	AC126591 Rattus no
32	231.6	24.8	720	6	AX365333	AX365333 Sequence
33	231.6	24.8	720	9	AF176117	AF176117 Homo sapi
34	230.2	24.6	184672	2	AC107850	AC107850 Mus muscu
35	221.6	23.7	169841	2	AC073888	AC073888 Homo sapi
36	221.6	23.7	199200	9	AL359633	AL359633 Human DNA
37	219.2	23.5	67395	2	AC134417	AC134417 Mus muscu
38	217	23.2	626	10	AB075973	AB075973 Rattus no
39	214.8	23.0	2216	10	CG060263	CG060263 Cricetus
40	212.2	22.7	621	10	AF282255	AF282255 Mus muscu
41	212.2	22.7	2125	10	AF260717	AF260717 Mus muscu
42	211	22.6	597	6	AX365312	AX365312 Sequence
43	211	22.6	597	10	AF237789	AF237789 Mus muscu
44	210.8	22.6	2198	10	BC013551	BC013551 Mus muscu
45	209.6	22.4	597	6	AX365315	AX365315 Sequence

ALIGNMENTS

RESULT 1
AX074351
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX074351
Sequence 3 from Patent WO0104305.
AX074351
AX074351.1 GI:12710514

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

Linear PAT 06-FEB-2001

Handwritten signature

Incyte Genomics, Inc. (US)	
Location/Qualification	1. .934
FEATURES	
source	

BASE COUNT	247 a	244 c	206 g	237 c
ORIGIN				

Query Match	100.0%	Score 934;	DB 6;	Length 934;
Best Local Similarity	100.0%;	Pred. No. 1.3e-285;	Indels	0;
Matches 934;	Conservative	0;	Mismatches	0;
QY	1	AGAATCATCTCATCGAATGAGAGCCACAGGCGACAGGACACAGTCCCTGAAGATGG	60	
DB	1	AGAATCATCTTCATCGGAATGAGAGCCACAGGCGACAGGACACAGTCCCTGAAGATGG	60	
QY	61	AGGACTTTTCTTACTGTGCTGATAGACAGGACTGGGCTGTCACTCGTTGTTTTCGAGA	120	
DB	61	AGGACTTTTCTTACTGTGCTGCATAGACAGGACTGGGCTGTCACTCGTTGTTTTCGAGA	120	
QY	121	AGAAGCCTTTCAGCAATCACATGACTTCAATGACCTCCCACTCGTTTTCGCGTCAA	180	
DB	121	AGAAGCCTTTCAGCAATCACATGACTTCAATGACCTCCCACTCGTTTTCGCGTCAA	180	
QY	181	TGTTCCACGACTAGTGTGGAAGGAGAGACGAAGAAATTTGAGGACATGTTTCG	240	
DB	181	TGTTCCACGACTAGTGTGGAAGGAGAGACGAAGAAATTTGAGGACATGTTTCG	240	
QY	241	GACTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGCTCGTATAAA	300	
DB	241	GACTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGCTCGTATAAA	300	
QY	301	CTTCAGCAATCCTAAATCTCCACCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAG	360	
DB	301	CTTCAGCAATCCTAAATCTCCACCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAG	360	
QY	361	AGGGAAAAAATTTAAAGCTCTACTTTGACACAGGTTTCAGACTCCAGACAGATGGAGACAA	420	
DB	361	AGGGAAAAAATTTAAAGCTCTACTTTGACACAGGTTTCAGACTCCAGACAGATGGAGACAA	420	
QY	421	ACTGCACATGGCTCCACCCGAGCTGCCAACAGTTTCTCATCTCGCCCCCTTCTCTCCC	480	
DB	421	ACTGCACATGGCTCCACCCGAGCTGCCAACAGTTTCTCATCTCGCCCCCTTCTCTCCC	480	
QY	481	ACCTCTGGCTGGCAGGCCATCAACGATGCCACGCCAGTCCCTCACTATGACCTCTCTA	540	
DB	481	ACCTCTGGCTGGCAGGCCATCAACGATGCCACGCCAGTCCCTCACTATGACCTCTCTA	540	
QY	541	TGCTGTGGCCAACTAGACACAGGAGAGAATGATGAGCTCCATCGAGGACTCAGTCCAC	600	
DB	541	TGCTGTGGCCAACTAGGACACAGGAGAGAATGATGAGCTCCATCGAGGACTCAGTCCAC	600	
QY	601	CCCAAGTCTGCTGTCAGTGTGCGACAGTGACATAGAGGAAGAGGACCCCAAGAC	660	
DB	601	CCCAAGTCTGCTGTCAGTGTGCGACAGTGACATAGAGGAAGAGGACCCCAAGAC	660	
QY	661	TTCCCAAGGCCAAAAATCATCCAACTCGGGTCTCGCCAGCCCTCCGTGTCCAA	720	
DB	661	TTCCCAAGGCCAAAAATCATCCAACTCGGGTCTCGCCAGCCCTCCGTGTCCAA	720	
QY	721	CTGAGCTGCTGCTCTTCGATATAGCCGCTCTCTCTTATCATGCTTTTTCGCCCT	780	
DB	721	CTGAGCTGCTGCTCTTCGATATAGCCGCTCTCTCTTATCATGCTTTTTCGCCCT	780	
QY	781	GTTCTTTGTCAAAAAAATTCGCTTTAAATTCCTGGTGTGTTGGTTTGGATTCCTT	840	
DB	781	GTTCTTTGTCAAAAAAATTCGCTTTAAATTCCTGGTGTGTTGGTTTGGATTCCTT	840	
QY	841	CCTTGTATTACGCTCTCGACAAAGGGCTAGGAAAAAGGTGATATGTCCTGATCAT	900	
DB	841	CCTTGTATTACGCTCTCGGACAAAGGGCTAGGAAAAAGGTGATATGTCCTGATCAT	900	

RESULT 5	
AY034086	LOCUS
AY034086	DEFINITION
AY034086	ACCESSION
AY034086	VERSION
AY034086.1	KEYWORDS
GI:21307624	SOURCE
Human sapiens (human)	ORGANISM
Human sapiens	REFERENCE
Eumalacostraca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	AUTHORS
1 (bases 1 to 3253)	TITLE
Cao,X., Kambe,F., Miyazaki,T., Sarkar,D., Ohmori,S. and Seo,H.	JOURNAL
Novel human ZAKI-4 isoforms: hormonal and tissue-specific regulation and function as calcineurin inhibitors	MEDLINE
Biochem. J. 367 (pt 2), 459-466 (2002)	PUBMED
22250686	REFERENCE
12102656	AUTHORS
2 (bases 1 to 3253)	TITLE
Cao,X., Kambe,F., Miyazaki,T., Ohmori,S. and Seo,H.	JOURNAL
Direct Submission	FEATURES
Submitted (08-MAY-2001) Endocrinology and Metabolism, Research Institute of Environmental Medicine, Nagoya University, Furo-cho, Nagoya 464-8601, Japan	
Location/Qualifiers	

Source

```

source
1. .3253
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p12"
/tissue_type="brain"
121..852
/notes="alternatively spliced"
/codon_start=1
/product="calcineurin inhibitor ZAKI-4 beta splice variant
2"
/protein_id="AAK59806.1"
/db_xref="GI:21307625"
/translation="MGESYFTGMRSPQOQHVEDGGFLFLCCIDRDWVTCFAEE
AFQAITDNDPLNSLFACHNVHVSFGEESKEKGLFRVYDDCVYFOLFKEFFRRVRI
NFSNPSAAKARLELHEHETQFRGKLKLYFAOVQTPETDGDKLHLAPPOPAKQLFLISPP
SSPVSQNPINDIATPVNLYDLVAVAKLGPGEKVELHAGTSESTPSVVVHVHCDSDIEEE
EDPRTSPKPKIQTTRPGLPPSVSN"
misc_feature
346..3253
/notes="similar to sequence with GenBank Accession Number
DB3407"
polyA_signal
1092..1097
/notes="alternative"
polyA_signal
3237..3242
BASE COUNT 938 a 707 c 683 g 925 t
ORIGIN
Query Match 99.7%; Score 930.8; DB 9; Length 3253;
Best Local Similarity 99.8%; Pred. No. 1.7e-284;
Matches 932; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGAATCATACTTCATCGGAATGAGAGCCAGCGAGGAGGACACGCTCCCTGGAAGATGG 60
Db |||||
129 AGAATCATACTTCATCGGAATGAGAGCCAGCGAGGAGGACACGCTCCCTGGAAGATGG 188
QY 61 AGGACTTTTCTTACTGTGCTGCATAGACAGGGACTGGGCTGTCACTCGTGTGTTGTCGAG 120
Db |||||
189 AGGACTTTTCTTACTGTGCTGCATAGACAGGGACTGGGCTGTCACTCGTGTGTTGTCGAG 248
QY 121 AGAAGCCTTTCAGCAATCACTGACTTCAATGACCTCCCACTCGTGTGTTGTCGAG 180
Db |||||
249 AGAAGCCTTTCAGCAATCACTGACTTCAATGACCTCCCACTCGTGTGTTGTCGAG 308
QY 181 TGTTCACAGTCAGTGTGTTGAAGGAGAGAGAGCAAGAAAATTTGAGGAGCTGTTTCG 240
Db |||||
309 TGTTCACAGTCAGTGTGTTGAAGGAGAGAGAGCAAGAAAATTTGAGGAGCTGTTTCG 368
QY 241 GACTTATGATGACTGTGTGACGTTCACGTATTTAAGAGTTTCAGAGCTGTCGATATAA 300
Db |||||
369 GACTTATGATGACTGTGTGACGTTCACGTATTTAAGAGTTTCAGAGCTGTCGATATAA 428
QY 301 CTTGAGCAATCCTTAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAAACCAATTTCAG 360
Db |||||
429 CTTGAGCAATCCTTAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAAACCAATTTCAG 488
QY 361 AGGGAAAAATTTAAAGCTCTACTTTGACAGGTTTCAGACTCCAGACAGATGGAGACAA 420
Db |||||
489 AGGGAAAAATTTAAAGCTCTACTTTGACAGGTTTCAGACTCCAGACAGATGGAGACAA 548
QY 421 ACTGCACTTGGCTCCACCCGAGCTGCCAACAGTTTCTCATCTCGCCCTTCTCTCC 480
Db |||||
549 ACTGCACTTGGCTCCACCCGAGCTGCCAACAGTTTCTCATCTCGCCCTTCTCTCC 608
QY 481 ACCTGTTGGCTGGCAGCCCATCAACATGCGAGCCAGTCTCAACTATGACCTCCTCTA 540
Db |||||
609 ACCTGTTAGCTGGCAGCCCATCAACATGCGAGCCAGTCTCAACTATGACCTCCTCTA 668
QY 541 TGCTGTGCCAACTAGGACCCAGGAGAGTATGAGCTCCATGCGAGGACTGAGTCCAC 600
Db |||||
669 TGCTGTGCCAACTAGGACCCAGGAGAGTATGAGCTCCATGCGAGGACTGAGTCCAC 728
QY 601 CCCAAGTGTGCTGTGCAGCTGTGCACAGTGACATAGAGAAAGAGGACCAAGAC 660

```

Tue Aug 26 09:11:04 2003

```

/product="Similar to Down syndrome critical region gene
1-like 1"
/protein_id="AAH38509.1"
/db_xref="GI:23620412"
/translation="MPAPSMDCOVSTLVACVVDVEFTNOEVEKFEGLFRYYDDCVT
FOLFKFRRVRINFSNPKSAARADIELHETQGRKKLKYFAQVOTPTDGDKLHLP
POPAKOFLLSPSPVPWPINDATFVLNLDLYAKLGGKYEYHACTETPSV
VHVCDSDIEEDDPKTSPIKLIQTRPLPSPVSNNAACSFSLIAYSSLSFCFPL
FVKKNCL"
BASE COUNT 755 a 559 c 501 g 708 t
ORIGIN
Query Match 77.2%; Score 720.8; DB 9; Length 2523;
Best Local Similarity 95.9%; Pred. No. 1.3e-217; Indels 0; Gaps 0;
Matches 740; Conservative 0; Mismatches 32;
163 CTCGTTGTTGGTGCATGTTCCACAGTCAGTGTGTTGAAGGAGAGAGCAAGGAAA 222
229 CACTCTGGTGGCTGTGGTGGATGTCGAGGTCTTTACCAATCAGGAGGTTAAGGAAA 288
223 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGACGTTCCAGCTATTAAAGAGTT 282
289 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGACGTTCCAGCTATTAAAGAGTT 348
283 CAGAGCTGTCCTGATATAAATTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGCT 342
349 CAGAGCTGTCCTGATATAAATTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGCT 408
343 TCATGAACCCAAATTCAGAGGAGGAAAATTAAGCTCTACCTTTCACAGAGTTTCAGACTCC 402
409 TCATGAACCCAAATTCAGAGGAGGAAAATTAAGCTCTACCTTTCACAGAGTTTCAGACTCC 468
403 AGACAGAGTGGAGACAAATGCACCTTGGCTCCACCCAGCCATCAAGATGCCACGAGTCT 462
469 AGACAGAGTGGAGACAAATGCACCTTGGCTCCACCCAGCCATCAAGATGCCACGAGTCT 528
463 CTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 522
529 CTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 588
523 CAACCTATGACCTCTCTATGCTGGCCAAATAGGACGAGGAGAGATGAGCTCCA 582
589 CAACCTATGACCTCTCTATGCTGGCCAAATAGGACGAGGAGAGATGAGCTCCA 648
583 TCGAGGAGTCTGAGTCCACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
649 TCGAGGAGTCTGAGTCCACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
643 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAACTCGGCTCTGCGCT 702
709 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAACTCGGCTCTGCGCT 768
703 GCACCTCCGCTGCTCAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762
769 GCACCTCCGCTGCTCAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
763 ATCATGCTTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 822
829 ATCATGCTTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 888
823 GGTGTTGAGATTCCTTCTCTTATCAAGCTCTCGGACAAAAGGCTAGGAAAAGGT 882
889 GGTGTTGAGATTCCTTCTCTTATCAAGCTCTCGGACAAAAGGCTAGGAAAAGGT 948
883 GATATGCTCTGATCATATATCAATACCATTAAGTATACCCATTATTTAGAA 934
949 GATATGCTCTGATCATATATCAATACCATTAAGTATACCCATTATTTAGAA 1000

```

RESULT 7
AX329596 3184 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 105 from Patent WO0194629.

```

AX329596
VERSION AX329596.1 GI:18102574
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 Young, P.E., Augustus M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 105 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1..3184
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 921 a 681 c 657 g 925 t
ORIGIN
Query Match 77.2%; Score 720.8; DB 6; Length 3184;
Best Local Similarity 95.9%; Pred. No. 1.4e-217; Indels 0; Gaps 0;
Matches 740; Conservative 0; Mismatches 32;
163 CTCGTTGTTGGTGCATGTTCCACAGTCAGTGTGTTGAAGGAGAGAGCAAGGAAA 222
222 CACTCTGGTGGCTGTGGTGGATGTCGAGGTCTTTACCAATCAGGAGGTTAAGGAAA 281
223 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGACGTTCCAGCTATTAAAGAGTT 282
282 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGACGTTCCAGCTATTAAAGAGTT 341
283 CAGAGCTGTCCTGATATAAATTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGCT 342
342 CAGAGCTGTCCTGATATAAATTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGCT 401
343 TCATGAACCCAAATTCAGAGGAGGAAAATTAAGCTCTACCTTTCACAGAGTTTCAGACTCC 402
402 TCATGAACCCAAATTCAGAGGAGGAAAATTAAGCTCTACCTTTCACAGAGTTTCAGACTCC 461
403 AGACAGAGTGGAGACAAATGCACCTTGGCTCCACCCAGCCATCAAGATGCCACGAGTCT 462
462 AGACAGAGTGGAGACAAATGCACCTTGGCTCCACCCAGCCATCAAGATGCCACGAGTCT 521
463 CTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 522
522 CTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 581
523 CAACCTATGACCTCTCTATGCTGGCCAAATAGGACGAGGAGAGATGAGCTCCA 582
582 CAACCTATGACCTCTCTATGCTGGCCAAATAGGACGAGGAGAGATGAGCTCCA 641
583 TCGAGGAGTCTGAGTCCACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
642 TCGAGGAGTCTGAGTCCACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
643 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAACTCGGCTCTGCGCT 702
702 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAACTCGGCTCTGCGCT 761
703 GCACCTCCGCTGCTCAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762
762 GCACCTCCGCTGCTCAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821
763 ATCATGCTTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 822
822 ATCATGCTTTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 881
823 GGTGTTGAGATTCCTTCTCTTATCAAGCTCTCGGACAAAAGGCTAGGAAAAGGT 882
882 GGTGTTGAGATTCCTTCTCTTATCAAGCTCTCGGACAAAAGGCTAGGAAAAGGT 941

```

QY 883 GATATGTCCTCATATCATATACCCATTAAGTATATAGCCATTAATTTAGAA 934
|||||
Db 942 GATATGTCCTCATATCATATACCCATTAAGTATATAGCCATTAATTTAGAA 993
|||||

RESULT 8
AX333015
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX333015
Sequence 3524 from Patent WO0194629.
AX333015
AX333015.1 GI:18123649
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R., and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 3524 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers
1..3184
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 921 a 681 c 657 g 925 t
ORIGIN

Query Match 77.2%; Score 720.8; DB 6; Length 3184;
Best Local Similarity 95.9%; Pred. No. 1.4e-217;
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 163 CTCGCTGTTTCGGTCATGTTCCACAGTCAGTGTGTTGAAGGAGAGAGAGAGAA 222
|||||
Db 222 CACTCTGTTGCCCTGTTGGTGGATGTCGAGGCTTTTACCAATCAGAGGTTAAGGAAA 281
|||||

QY 223 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAGAGTTT 282
|||||
Db 282 ATTTGGGGAGCTGTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAGAGTTT 341
|||||

QY 283 CAGACGTCCTGTTAACTTCAGCAATCTTAATCTGACGCCGAGCTAGGATAGACT 342
|||||
Db 342 CAGACGTCCTGTTAACTTCAGCAATCTTAATCTGACGCCGAGCTAGGATAGACT 401
|||||

QY 343 TCATGAAACCAATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
|||||
Db 402 TCATGAAACCAATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
|||||

QY 403 AGACAGATGGAG 462
|||||
Db 462 AGACAGATGGAG 521
|||||

QY 463 CTCGCCCTTCCTCCCTCCCTGTTGGCTGGAGCCATCAAGATGCCAGCCAGTCT 522
|||||
Db 522 CTCGCCCTTCCTCCCTCCCTGTTGGCTGGAGCCATCAAGATGCCAGCCAGTCT 581
|||||

QY 523 CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACAGGAGAGAGAGAGAGAG 582
|||||
Db 582 CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACAGGAGAGAGAGAGAGAG 641
|||||

QY 583 TGCAGGAGCTGAGTCCACCCCAAGTCTGCTGTCGACGTGTGCGACAGTACATAGAGGA 642
|||||
Db 642 TGCAGGAGCTGAGTCCACCCCAAGTCTGCTGTCGACGTGTGCGACAGTACATAGAGGA 701
|||||

QY 643 AGAAGAGGAG 702
|||||
Db 702 AGAAGAGGAG 761
|||||

QY 703 GCCACCTCCGTCGTCAGTGTGCTGCTCTCTTCGATATAGCCGCTCTCTCTTT 762
|||||

Db 762 GCCACCTCCGTCGTCAGTGTGCTCTCTTCGATATAGCCGCTCTCTCTTT 821
|||||

QY 763 ATCATGCTTTTCCCTCTGTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTT 822
|||||

Db 822 ATCATGCTTTTCCCTCTGTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTT 881
|||||

QY 823 GPTGTTTTCAGATTCCTCTGTTTATCAAGCTCTCGGACAAAAGGCTAGGAAAGGT 882
|||||

Db 882 GPTGTTTTCAGATTCCTCTGTTTATCAAGCTCTCGGACAAAAGGCTAGGAAAGGT 941
|||||

QY 883 GATATGTCCTCATATCATATACCCATTAAGTATATAGCCATTAATTTAGAA 934
|||||

Db 942 GATATGTCCTCATATCATATACCCATTAAGTATATAGCCATTAATTTAGAA 993
|||||

RESULT 9
AX365327
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX365327
Sequence 17 from Patent WO0204491.
AX365327
AX365327.1 GI:18697053
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Williams, S.R. and Rothermel, B.
Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)
Patent: WO 0204491-A 17 17-JAN-2002;
Board of Regents, The University of Texas System (US); Williams,
Sanders R. (US); Rothermel, Beverly (US)
Location/Qualifiers
1..3184
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
205..783
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23813.1"
/db_xref="GI:18697054"
/translation="MDCDVSTLVACVYDVEFTNOEVKEKFGLEFRTYDDCVTFOLFK
SFRVRINFSNPKSAARARIELHETQFRGKILKLYFAOVOTPETDGLHLAPQPAK
QFLISPPSPVSWQPINADLVNLDLYAVAKLPGKEKELHAGTETPSVVVHVC
DSDIEEDPKTSPKPKIIQTRPGLPSPVSN"

BASE COUNT 921 a 681 c 657 g 925 t
ORIGIN

Query Match 77.2%; Score 720.8; DB 6; Length 3184;
Best Local Similarity 95.9%; Pred. No. 1.4e-217;
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 163 CTCGCTGTTTCGGTCATGTTCCACAGTCAGTGTGTTGAAGGAGAGAGAGAGAA 222
|||||
Db 222 CACTCTGTTGCCCTGTTGGTGGATGTCGAGGCTTTTACCAATCAGAGGTTAAGGAAA 281
|||||

QY 223 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAGAGTTT 282
|||||
Db 282 ATTTGGGGAGCTGTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAGAGTTT 341
|||||

QY 283 CAGACGTCCTGTTAACTTCAGCAATCTTAATCTGACGCCGAGCTAGGATAGACT 342
|||||
Db 342 CAGACGTCCTGTTAACTTCAGCAATCTTAATCTGACGCCGAGCTAGGATAGACT 401
|||||

QY 343 TCATGAAACCAATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
|||||
Db 402 TCATGAAACCAATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
|||||

QY 403 AGACAGATGGAG 462
|||||

Query Match					
Best Local Similarity		Score 720.8; DB 9;	Length 3184;		
Matches 740; Conservative		Pred. No. 1.4e-217;			
		0; Mismatches 32;	Indels	0;	Gaps
QY	163	CTCGTGTGTTGGTCGCATCATGTTCCACCAAGTCTTGTAAGGACAGACAGCAGCAAGAAAA	222		
Dd	222	CACCTCGTGTGCTGTGTGGTGAATGTCGAGGTCTTTACCAATCAGAGGTTAAAGAAAA	281		
QY	223	ATTTCGAGGACTGTTTCGGACTATGATGACTGTGTGACGTTCACGCTATTTAAAGAGTTT	282		
Dd	282	ATTTCGGGCACTGTTTCGGACTTATGATGACTGTGTGACGTTCACGCTATTTAAAGAGTTT	341		
QY	283	CAGACGTGTCGGTATAAACHTTCAGCAATCCTAAATCTGCAGCCCAGCTAGATAGAGCT	342		
Dd	342	CAGACGTGTCGTATAAACHTTCAGCAATCCTAAATCTGCAGCCCAGCTAGATAGAGCT	401		
QY	343	TCATGAACCCAATTTCAGAGGGAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCC	402		
Dd	402	TCATGAACCCAATTTCAGAGGGAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCC	461		
QY	403	AGACACAGATGGAGACAACACTGCACCTTGGCTCCACCAGGCTGCCAAACAGTTTTCAT	462		
Dd	462	AGACACAGATGGAGACAACACTGCACCTTGGCTCCACCAGGCTGCCAAACAGTTTTCAT	521		
QY	463	CTCGCCCCCTTCCTCCCAACCTGTTGGCTGGCAGCCCATCAACCATGCCAGCTCCT	522		
Dd	522	CTCGCCCCCTTCCTCCCAACCTGTTAGCTGGCAGCCCATCAACCATGCCAGCTCCT	581		
QY	523	CAACTATGACCTTCCTATGCTGTGGCCAACTAGCACAGGAGAAGATATGAGTCCA	582		
Dd	582	CAACTATGACCTTCCTATGCTGTGGCCAACTAGCACAGGAGAAGATATGAGTCCA	641		
QY	583	TGCAGGGAATGAGTCCACCCCAAGTGTGCTGCACGTGTGGACAGTGCATATGAGGA	642		
Dd	642	TGCAGGGAATGAGTCCACCCCAAGTGTGCTGCACGTGTGGACAGTGCATATGAGGA	701		
QY	643	AGAAGAGACCAAGAGACTTCCCAAGCCAAAATCATCAAATCGCGCTCTGGCT	702		
Dd	702	AGAAGAGACCAAGAGACTTCCCAAGCCAAAATCATCAAATCGCGCTCTGGCT	761		
QY	703	GCCACCTCGGTGCAACTGAGCTGCCTTCTCGATAAATAGCCGCTCTCTTT	762		
Dd	762	GCCACCTCGGTGCAACTGAGCTGCCTTCTCGATAAATAGCCGCTCTCTTT	821		

Db 739 TTCCCAAGCCAAAATCATCCAGACCCCGCGTCTCGCTGCCACCCCTCCGTCCTCAA 798
 QY 721 CTGAGCTCCCTGCTCTCTCGATAATAGCGCTCTCTCTTTATCATGCTTTTCCCTCC 780
 Db 799 CTGA-----GCACGTCCTATGATAACATCATCTCTCTC-TTATCATGCTTTTCTCCCT 850
 QY 781 GTTCTTTGTCACAAAATTCCTTTAAATTCCTGGGTGTTGGTGTGTTGAGATTCCTT 840
 Db 851 TGTTCGTCCAAAATTAATTCCTTGAAGTCCCTGGGTGTTGGTGTGTTGAGATTCCTT 910
 QY 841 CTTTGTATCAAGCTCTCGGACAAAAGGCTAGGAAAAGGTG 883
 Db 911 CTTGTATTAAGCTCTCGGACAAAAGGCTAGGAAAAGGTG 953

RESULT 14

BC047153

LOCUS

DEFINITION

Mus musculus, similar to down syndrome critical region gene 1-like

ACCESSION

BC047153

VERSION

BC047153.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: comphs-femail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Avello, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granito, S., Guan, X., Gupta, J., Hagnighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 100 Row: n Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13447391.

FEATURES

source

1. 3080
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6491318"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH_MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 882 a 677 c 678 g 843 t

Query Match

55.1%; Score 514.6; DB 10; Length 3080;

Best Local Similarity 79.9%; Pred. No. 6.6e-152;
 Matches 618; Conservative 0; Mismatches 154; Indels 1; Gaps 1;

QY 163 CTCCTGTTGTCGCTGCAATGTTCCACAGTCAGTGTGTTGAAGGAGAAGAGAGCAAGGAAA 222
 Db 91 CACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150
 QY 223 ATTGAGGAGTCTGTTCCGGACCTTATGATGACTGTGTGACGTTTCCAGCTATTTAAAGATT 282
 Db 151 ATTCGAGGAGTCTGTTCCGGACCTTATGATGAAATGTGTGACGTTTCCAGCTGTTTAAAGATT 210
 QY 283 CACAGCTGTCCTCGTATAAATTCAGCAATCTTAAATCTGACGCTCCAGCCGCTAGATAGAGCT 342
 Db 211 CGACGGGTTCGAATAAATTCAGCCATCCCAATCTGACGCCGCTGCGGATAGAGCT 270
 QY 343 TCATGAACCCCAATTCAGAGGGAATAAATTAAGCTCTACTTTGCACAGTTCAGAGTCC 402
 Db 271 TCATGAGACTCAGTTCAGAGGGAATAAATTCAGCTTCTACTTCCGCGAGTCCAGACCCC 330
 QY 403 AGAGACAGATGAGAGCAAACTGCACTTGGCTCCACCCAGCTCCCAACAGTTCCTCAT 462
 Db 331 AGAGACAGATGAGAGCAAACTGCACTTGGCTCCACCCAGCTCCCAACAGTTCCTCAT 390
 QY 463 CTCGCCCT 522
 Db 391 CTCGCCCT 450
 QY 523 CAACATGACCT 582
 Db 451 CAACATGACCT 510
 QY 583 TGCAGGACTGAGTCCACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
 Db 511 CGCTGGAACGTGAGTCTACCCGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
 QY 643 AGAAGAGGAGCCCAAGACCTTCCCAAGCCAAATATCACTCCAACTCGCGCTCTCTCTCTCTCT 702
 Db 571 GGAGGAGGAGCCCAAGACCTTCCCAAGCCAAATATCACTCCAGACCCGCTCTCTCTCTCTCT 630
 QY 703 GCCACCTCCGCTGCTCAACTGAGC-TGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761
 Db 631 GCCACCTCCGCTGCTCAACTGAGCAGCTCATGATGATGATGATGATGATGATGATGATGATGAT 690
 QY 762 TATCATGCTTTTCCCT 821
 Db 691 TCCCGCT 750
 QY 822 TGGTGTGTTGAGATTCCT 881
 Db 751 TGGTGTGTTGAGATTCCT 810
 QY 882 TGATATGCTCTCTGATCATATCATACCATTAAGTATAACCCATTTATTAGAA 934
 Db 811 TGATATGCTCTCTCAATCAACTCTTACCCATTAACATCACTCACTTAAAG 863

RESULT 15

BC049096

LOCUS

DEFINITION

Mus musculus, Down syndrome critical region gene 1-like 1, clone

ACCESSION

BC049096

VERSION

BC049096.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

BC049096

Mus musculus, Down syndrome critical region gene 1-like 1, clone

MGC:61350 IMAGE:6416901, mRNA, complete cds.

BC049096.1

GI:29145001

MGC.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Strausberg, R.

Direct Submission

Submitted (19-MAR-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

||||| 653 CAACTATGACCTTCTTTATGCTGTGGCCAACTAGGACGAGAGAAATATGAGCTGCA 712
||||| 583 TCGAGGAGTACTGAGTCCACCCCAAGTGTGCTGTGACGCTGCGACAGTGCATAGAGGA 642
||||| 713 CGCTGGAACCTGAGTCTACACCGAGCGTGTGTGATGTGTGACAGGACATGAGGA 772
||||| 643 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAAACTCGGCGCTCTGGGCT 702
||||| 773 GGAGGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAAACTCGGCGCTCTGGGCT 832
||||| 703 GCCACCTCGGTGTCCAACTGAGC-TGCCCTGCTCTCGATATAGCCGCTCTCTCT 761
||||| 833 GCCACCTCGGTGTCCAACTGAGCAGCTCCATGATACATGCTCTCTCTCTCTCT 892
||||| 762 TATCATGCTTTTTCCTGCTGTTTGTCAAAAAAATTCCTTTTAAATTCCTTGGGTGT 821
||||| 893 TCCGCGCCCGCCCTTGTGTTGTCCTCAAAATTAATTTGCTTTAAGTCCCTGGGTGT 952
||||| 822 TGGTGTGTTGAGATTCCTTCTTGTATCAAGCTCTCGGACAAAGGCTAGGAAAGG 881
||||| 953 TGGTGTGTTGAGATTCCTTCTTGTATCAAGCTCTCGGACAAAGGCTAGGAAAGG 1012
||||| 882 TGATATGCTCTCTGATCATATCATATCAATCAATCAATCAATCAATCAATCAAT 934
||||| 1013 TGATATGCTCTCTGATCATATCATATCAATCAATCAATCAATCAATCAATCAAT 1065

Search completed: August 24, 2003, 11:27:03
Job time : 4686 secs

Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY
653 CAACTATGACCTTCTTTATGCTGTGGCCAACTAGGACGAGAGAAATATGAGCTGCA 712
583 TCGAGGAGTACTGAGTCCACCCCAAGTGTGCTGTGACGCTGCGACAGTGCATAGAGGA 642
713 CGCTGGAACCTGAGTCTACACCGAGCGTGTGTGATGTGTGACAGGACATGAGGA 772
643 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAAACTCGGCGCTCTGGGCT 702
773 GGAGGAGGACCCAAAGACTTCCCAAGCCAAATCATCCAAACTCGGCGCTCTGGGCT 832
703 GCCACCTCGGTGTCCAACTGAGC-TGCCCTGCTCTCGATATAGCCGCTCTCTCT 761
833 GCCACCTCGGTGTCCAACTGAGCAGCTCCATGATACATGCTCTCTCTCTCTCT 892
762 TATCATGCTTTTTCCTGCTGTTTGTCAAAAAAATTCCTTTTAAATTCCTTGGGTGT 821
893 TCCGCGCCCGCCCTTGTGTTGTCCTCAAAATTAATTTGCTTTAAGTCCCTGGGTGT 952
822 TGGTGTGTTGAGATTCCTTCTTGTATCAAGCTCTCGGACAAAGGCTAGGAAAGG 881
953 TGGTGTGTTGAGATTCCTTCTTGTATCAAGCTCTCGGACAAAGGCTAGGAAAGG 1012
882 TGATATGCTCTCTGATCATATCATATCAATCAATCAATCAATCAATCAATCAAT 934
1013 TGATATGCTCTCTGATCATATCATATCAATCAATCAATCAATCAATCAATCAAT 1065

REMARK
COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: craps-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,
Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
Scheetz, J., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,
Casavant, T., Soares, M.B.

FEATURES
source

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0.
Location/Qualifiers
1. 3263
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:61350 IMAGE:6416901"
/tissue_type="Mouse, brain 12.5 dpc"
/clone_lib="NIH_BMAP_F0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
261..854
/codon_start=1
/product="down syndrome critical region gene 1-like 1"
/protein_id="AAH49096.1"
/db_xref="GI:29145002"
/db_xref="LocusID:53901"
/translation="MPAPSMDCDVSTLVCVVDVEFTNOEVEKPEGLRTYDECVT
POLKSPRRVRFNHPKSAARARIEHETOFGRKLUKLYFAQVQVPEPTDGLHLAP
POLAKOFLISPPSPVWPKPISDATVLYDLYAVAKLGPGEKELHAGTSTPSV
VHVCDSDMEEDPKTSPKPIITRRPGLPSPVSN"
BASE COUNT 907 a 732 c 748 g 876 t

CDS

Query Match 55.1%; Score 514.6; DB 10; Length 3263;
Best Local Similarity 79.9%; Pred. No. 6.6e-152; Indels 1; Gaps 1;
Matches 618; Conservative
163 CTCGTGTTGCTGCAATGTTCCACAGTCAGTGTGTTGAAGGAGAGAGAGAGAGAGAG 222
163 CTCGTGTTGCTGCAATGTTCCACAGTCAGTGTGTTGAAGGAGAGAGAGAGAGAGAG 222
293 CACTCTGGTGGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 352
293 CACTCTGGTGGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 352
223 ATTTGAGGAGTCTTTTCGGACTATGATGCTGTGTGAGTTCAGCTATTTAAAGAGTTT 282
223 ATTTGAGGAGTCTTTTCGGACTATGATGCTGTGTGAGTTCAGCTATTTAAAGAGTTT 282
353 ATTCGAGGAGTCTTCCGGACCTATGATGATGATGATGATGATGATGATGATGATGAT 412
353 ATTCGAGGAGTCTTCCGGACCTATGATGATGATGATGATGATGATGATGATGATGAT 412
283 CAGACGTGCTCGGTATTAACCTCAGCAATCTCAATCTCAATCTCAATCTCAATCTCA 342
283 CAGACGTGCTCGGTATTAACCTCAGCAATCTCAATCTCAATCTCAATCTCAATCTCA 342
413 CCGACGGGTTCGAATTAATTTAGCCATCCCAATCTCGAGCCGCTGCGCGGATAGAGCT 472
413 CCGACGGGTTCGAATTAATTTAGCCATCCCAATCTCGAGCCGCTGCGCGGATAGAGCT 472
343 TCATGAACCCCAATTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
343 TCATGAACCCCAATTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
473 TCATGAGACTCAGTTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532
473 TCATGAGACTCAGTTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532
403 AGAGACAGATGAGACAAACTGACATTTGGTTCACACCCAGCCCTGCCAAACAGTTTCT 462
403 AGAGACAGATGAGACAAACTGACATTTGGTTCACACCCAGCCCTGCCAAACAGTTTCT 462
533 AGAGACAGATGAGACAAACTGACATTTGGTTCACACCCAGCCCTGCCAAACAGTTTCT 592
533 AGAGACAGATGAGACAAACTGACATTTGGTTCACACCCAGCCCTGCCAAACAGTTTCT 592
463 CTCGCCCTTCTCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 522
463 CTCGCCCTTCTCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 522
593 CTCACCCCTTCT 652
593 CTCACCCCTTCT 652
523 CAACATATGAGCT 582
523 CAACATATGAGCT 582

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 18:54:42 ; Search time 29 Seconds
(without alignments)
341.405 Million cell updates/sec

Title: US-10-030-613-1
Perfect score: 1254
Sequence: 1 MRSFGQGHVPEGGFLC.....SPKPIQTRRGLPPSVSN 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	255	4	US-09-614-474-2
2	913	72.8	192	4	US-09-614-474-10
3	729.5	58.2	241	4	US-09-614-474-11
4	581	46.3	171	2	US-08-665-040-2
5	98	7.8	454	3	US-08-348-518C-4
6	98	7.8	454	3	US-08-476-509B-4
7	94	7.5	905	3	US-08-574-959A-9
8	94	7.5	905	3	US-09-357-014-9
9	94	7.5	1135	3	US-08-574-959A-7
10	94	7.5	1135	3	US-09-357-014-7
11	90	7.2	655	3	US-09-347-833-4
12	83.5	6.7	659	4	US-09-562-737-12
13	83	6.6	554	4	US-09-328-352-7796
14	82	6.5	243	1	US-07-960-112B-2
15	82	6.5	243	1	US-08-301-316B-2
16	82	6.5	243	3	US-08-473-399B-2
17	82	6.5	243	3	US-08-853-831-2
18	82	6.5	243	4	US-09-510-885-2
19	82	6.5	243	5	PCR-US93-09774-2
20	81.5	6.5	906	4	US-09-651-656-13
21	81.5	6.5	906	4	US-09-650-855-13
22	80.5	6.4	415	4	US-09-252-991A-31348
23	80.5	6.4	830	4	US-09-252-991A-22004
24	80	6.4	428	3	US-09-347-833-6
25	79.5	6.3	424	4	US-09-333-593A-8
26	79.5	6.3	1706	2	US-08-459-568-2
27	79.5	6.3	1706	2	US-08-399-411-2

28	79.5	6.3	1706	3	US-08-516-859A-2	Sequence 2, Appli
29	79.5	6.3	1706	4	US-09-586-472-2	Sequence 2, Appli
30	79.5	6.3	1706	4	US-09-528-706-2	Sequence 2, Appli
31	79	6.3	434	2	US-08-815-718-3	Sequence 3, Appli
32	79	6.3	529	4	US-09-252-991A-18830	Sequence 18830, A
33	79	6.3	529	2	US-08-884-072-6	Sequence 6, Appli
34	79	6.3	559	4	US-09-212-168-6	Sequence 6, Appli
35	78.5	6.3	580	4	US-09-252-991A-29484	Sequence 29484, A
36	77.5	6.2	350	2	US-08-960-022-8	Sequence 8, Appli
37	77.5	6.2	782	4	US-09-252-991A-30464	Sequence 30464, A
38	77	6.1	681	6	5194595-19	Patent No. 5194595
39	76.5	6.1	330	3	US-09-145-391-2	Sequence 2, Appli
40	76	6.1	566	4	US-09-252-991A-16790	Sequence 16790, A
41	76	6.1	714	4	US-09-347-878-16	Sequence 16, Appli
42	75	6.0	930	4	US-09-254-594-6	Sequence 6, Appli
43	75	6.0	2441	1	US-08-194-468-2	Sequence 2, Appli
44	75	6.0	2441	3	US-08-961-739-2	Sequence 2, Appli
45	75	6.0	2441	4	US-09-514-247A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-614-474-2
; Sequence 2, Application US/09614474
; Patent No. 6524819

GENERAL INFORMATION:
; APPLICANT: Jor Ing, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PG-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6524819 247500.5
; NAME/KEY: unsure
; LOCATION: 11
; OTHER INFORMATION: unknown or other
US-09-614-474-2

Query Match	100.0%	Score	1254;	DB	4;	Length	255;
Best Local Similarity	100.0%	Pred. No.	9.5e-137;				
Matches	234;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MRSFGQGHVPEGGFLC	DRDWA	TRCF	AEFAQ	ITD	FNLSL
Db	22	MRSFGQGHVPEGGFLC	DRDWA	TRCF	AEFAQ	ITD	FNLSL
QY	61	EGEESKEFGLRTYD	DCVTF	QFLK	FSRR	VRIN	FNFS
Db	82	EGEESKEFGLRTYD	DCVTF	QFLK	FSRR	VRIN	FNFS
QY	121	YFAQVQTPETD	GDGLH	LAP	QPA	KOFL	ISPP
Db	142	YFAQVQTPETD	GDGLH	LAP	QPA	KOFL	ISPP
QY	181	PGEKYLHAGT	ESTP	SVV	HHV	CDSD	IEED
Db	202	PGEKYLHAGT	ESTP	SVV	HHV	CDSD	IEED

RESULT 2
US-09-614-474-10

```
; Sequence 10, Application US/09614474
; Patent No. 6524819
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PC-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6524819 g1435040
US-09-614-474-10

Query Match 72.8%; Score 913; DB 4; Length 192;
Best Local Similarity 93.0%; Pred. No. 1.7e-97;
Matches 174; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 48 NSLFACNHQSVFEGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSPKSAARARIE 107
Db 6 SYLVACVVDVEFTNQEVKEFGGLFRTYDDCVTFQLFKSFRRVRINFSPKSAARARIE 65
QY 108 LHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKOFLISPPSSPPVGVQNPINDATPV 167
Db 66 LHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKOFLISPPSSPPVGVQNPINDATPV 125
QY 168 LNYDLLYAVAKLPGKGYELHAGTSTPSVVHVHVCDSIDIEEDPKTSPKPKIIQTRRP 227
Db 125 LNYDLLYAVAKLPGKGYELHAGTSTPSVVHVHVCDSIDIEEDPKTSPKPKIIQTRRP 185
QY 228 LPPSVSN 234
Db 186 LPPSVSN 192

RESULT 3
US-09-614-474-11
; Sequence 11, Application US/09614474
; Patent No. 6524819
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PC-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6524819 g6017919
US-09-614-474-11

Query Match 58.2%; Score 729.5; DB 4; Length 241;
Best Local Similarity 66.1%; Pred. No. 4.1e-76;
Matches 144; Conservative 22; Mismatches 49; Indels 3; Gaps 2;

QY 19 LCCIDRDWAVTRCEFA--EEAFQAITDFNDLPSLFACNVHOSVFEGESKEFGLERTY 76
Db 19 LCCIDRDWAVTRCEFA--EEAFQAITDFNDLPSLFACNVHOSVFEGESKEFGLERTY 76
```

```
Db 16 LCSTDQEEEMIFGENEDDLDEMMDLSLPTSLFACSVHEAVFEAREOKERPEALFTTY 75
QY 77 DDCVTFQLFKSFRRVRINFSPKSAARARIELHETQFRGKKLYFAQVQTPETDGDGKLH 136
Db 76 DDQVTFQLFKSFRRVRINFSPKSAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSY 135
QY 137 LAPQPAKQFLISPPSSPPVGVQNPINDATPVLYNDLLYAVAKLPGKGYELHAGTSTPS 196
Db 136 LLPPQPVKQFLISPPASPPVGVQMSQSDAMPVINYDLICAVSKLPGKGYELHAGTSTPS 195
QY 197 VVHVCDSDIEEDPKTSPKPKIIQTRRGLPPSVSN 234
Db 196 VVHVCDSETEEEETK-NPKQKIAQTRRPPPTAALN 232

RESULT 4
US-08-665-040-2
; Sequence 2, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUENTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61ST STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1 FOR DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,040
; FILING DATE: JUNE 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ES P9501140
; FILING DATE: JUNE 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: JANET I. CORD
; REGISTRATION NUMBER: 33,778
; REFERENCE/DOCKET NUMBER: U010815-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-665-040-2

Query Match 46.3%; Score 581; DB 2; Length 171;
Best Local Similarity 67.1%; Pred. No. 3.7e-59;
Matches 112; Conservative 22; Mismatches 25; Indels 8; Gaps 2;

QY 68 KFEGLFRTYDDCVTFQLFKSFRRVRINFSPKSAARARIELHETQFRGKKLYFAQVOT 127
Db 5 KFSLEFRTYDDITTFQYFKSKVRINFSPSAADARLQHLKTEFLGKEMKLYFAQTLH 64
```

QY	128	PETDGDKLLHAPQPAKQFLIPSSPPVQWQPIN	DATPVLNLDLYAVAKLPGGKEYEL	187
	:			
Db	65	IGSS----	HLAPPNDKFLISSPASPVGQVEDATPV	120
	:			
QY	188	HAGTSTSTSVVHVCDSDIEEDPKTS----	PKKIIQTRRGLPP	230
	:			
Db	121	HAATDTTSPVVHVHVCSDQEKKEEEMERMR	PKKIIQTRRPEYTP	167

RESULT 5

US-08-348-518C-4
Sequence 4, Application US/08348518C
Patent No. 6022740
GENERAL INFORMATION:
APPLICANT: SUDOL, MARIUS
APPLICANT: PEER, BORK
APPLICANT: HENRY, CHEN
TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

07/801
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,518C
FILING DATE: 01-DEC-1994
CLASSIFICATION: 514
ATTORNEY: 514

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-3800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 454 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
18-348-518C-4

Query Match	7.8%	Score 98;	DB 3;	Length 454;
Best Local Similarity	31.1%	Pred. No	0.012	
Matches	32			

Qy	139	PPQPAKFLISPPSSPPVGMQIPINDATPVNLDLLYAVAKLGPCKETELIAGTSTPSV- 197	5;
Db	8	PPQAPAGQGQSPQSPQPGQGPPS-----GGQ-----PAPAAATQAAQAP 48	
Qy	198	-----VVHV-CDSDIEEEE-----DPKTSKPKLIOTRRGLPPS 231	
Db	49	PAGHQVHVHRGDSGTDLEALFNVMNPKATNVQTPVPMRLKLPDS 94	

RESULT 6

US-08-476-509B-4
Sequence 4, Application US/08476509B
Patent No. 6034212
GENERAL INFORMATION:
APPLICANT: SUDOL, MARIUS

APPLICANT: PEER, BORK
 APPLICANT: HENRY, CHEN
 TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
 TITLE OF INVENTION: A SIGNALLING DOMAIN THEREIN, NUCLEIC
 TITLE OF INVENTION: PROTEIN AND THE DOMAIN, THEREOF
 TITLE OF INVENTION: THEREOF
 NUMBER OF INVENTION: 50
 NUMBER OF INVENTION: 50
 CORRESPONDENCE: 50

CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

ZIP: 07601
COMPUTER READABLE FORM

COMPUTER READABLE FORM: disk
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,509B
 FILING DATE: 01-DEC-1994
 CLASSIFICATION:

CLASSIFICATION:
ATTORNEY / RECENT

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-101 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684

TELEX: 133521
INFORMATION FOR CEE

```

; INFORMATION FOR SEQ ID NO:
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 454 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; US-08-476-509B-4

```

US-08-476-509B-4

Query Match	7.8%;	Score 98;	DB 3;	Length 454;
Best Local Similarity	31.1%;	Pred. No. 0.012;		
Matches 33;	Conservative 11;			

[illegible]

RESULT 7

US-08-574-959A-9
; Sequence 9, Application US/08574959A
; Patent No. 5962224

04224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SPACES: 1

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston

STATE: MASSACHUSETTS

COUNTRY: USA

ACCOUNT: USA
ZIP: 02109-1875

COMPUTER READABLE FILE

MEDIUM TYPE: F1000

COMPUTER: IBM PC compatible
FLOPPY disk

IBM PC compatible
OPERATING SYSTEM: PC-DOS

PC-DOS/MS-DOS

[illegible]

```

1  RESULT 8
2  US-09-357-014-9
3  ; Sequence 9, Application US/09357014
4  ; Patent No. 6291645
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
7  ; and Jack L. Strominger
8  ; TITLE OF INVENTION: PE2 POLYPEPTIDES, RELATED POLYPEPTIDES
9  ; AND USES THEREFOR
10 ;
11 ; NUMBER OF SEQUENCES: 22
12 ;
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: LAHIVE & COCKFIELD
15 ; STREET: 60 State Street, Suite 510
16 ; CITY: Boston
17 ; STATE: Massachusetts
18 ; COUNTRY: USA
19 ; ZIP: 02109-1875
20 ;
21 ; COMPUTER READABLE FORM:
22 ; MEDIUM TYPE: Floppy disk
23 ; COMPUTER: IBM PC compatible
24 ; OPERATING SYSTEM: PC-DOS/MS-DOS
25 ; SOFTWARE: Patent In Release #1.0, Version #1.25
26 ;
27 ; CURRENT APPLICATION NUMBER: US/09/357,014
28 ; FILING DATE: 19-Jul-1999
29 ;
30 ; PRIOR APPLICATION NUMBER: 08/574,959
31 ; APPLICATION DATA:
32 ; FILING DATE: <Unknown>
33 ;
34 ; ATTORNEY/AGENT INFORMATION:
35 ; NAME: Mandragouras, Amy E.
36 ; REGISTRATION NUMBER: 36,207
37 ; REFERENCE/DOCKET NUMBER: DFN-008
38 ;
39 ; TELECOMMUNICATION INFORMATION:
40 ; TELEPHONE: (617)227-7400
41 ; TELEFAX: (617)227-5941
42 ;
43 ; INFORMATION FOR SEQ ID NO: 9:
44 ; SEQUENCE CHARACTERISTICS:

```

[illegible]

RESULT 9
...-08-574-959A-7

001-574-959A-7
 US-08-574-959A-7 Application US/08574959A
 Patent No. 5962224
 GENERAL INFORMATION:
 APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadiamudi
 APPLICANT: and Jack L. Strominger
 TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
 AND USES THEREFOR
 TITLE OF INVENTION: AND USES THEREFOR
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, Suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/574,959A
 FILING DATE: 19-DEC-95
 ATTORNEY/AGENT INFORMATION:
 NAME: Mandragouras, Amy E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: DFN-008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1135 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 001-574-959A-7

```

JS-08-574-959A-7
Query Match          7.5%; Score 94; DB 2; Length 1135;
Best Local Similarity 23.7%; Pred: No. 0.14;
Matches 41; Conservative 20; Mismatches 80; Indels 32; Gaps 6;

61 EGEEKFGLFRTYDDCVTQLFKSFRVRVINFSPKSAARIELHETQFGKKKL 120
  |||..||| : : : ||| :
QY : : : : : EEEEEEEEEEEEEE- - - - - 966

928 EEEEEEEFEFEFEFEG - - - - - EEEEEEEEEEEEEE- - - - -
Db : : : : :

131 YFAQVQPTDGDKLHLAPQAPAKQFL19PPSSPPVGWQPINATPVLNLYAVAKLG 180
QY : : : : :

```


Db 967 ---DLEFGTAGVEGAPPTLPALPPESP-KVOPEPEP-----GILLEVEE--- 1016
Qy 181 PGKYLHAGTSTSPVHVHCDSDIEE-EDPKTSPKPKIITRRPGLPPSV 232
Db 1017 PGTEERGADTAPTLAEPALPSOGEVEREGESPAAGPPQQLVEEPEPSXPTL 1069

RESULT 10

US-09-357-014-7
; Sequence 7, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/357,014
; FILING DATE: 19-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/574,959
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 7.5%; Score 94; DB 3; Length 1135;
Best Local Similarity 23.7%; Pred. No. 0.14;
Matches 41; Conservative 20; Mismatches 80; Indels 32; Gaps 6;
Qy 61 EGESKEKEGLRTYDDCVTFQLFSAFRVRINFSNPKSAARIELHETQFGKLLK 120
Db 928 EEEEEEEEEEVEEG-----ELEEEEEEEEEEVEE----- 966
Qy 121 YFAQVQTPETDGLKHLAPPQAKQFLISPPSPVGVQWQIPNDATPVLYNLYAVAKLG 180
Db 967 ---DLEFGTAGVEGAPPTLPALPPESP-KVOPEPEP-----GILLEVEE--- 1016
Qy 181 PGKYLHAGTSTSPVHVHCDSDIEE-EDPKTSPKPKIITRRPGLPPSV 232
Db 1017 PGTEERGADTAPTLAEPALPSOGEVEREGESPAAGPPQQLVEEPEPSXPTL 1069

RESULT 11

US-09-347-833-4
; Sequence 4, Application US/09347833
; Patent No. 6294658

; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Odell, Joan T.
; TITLE OF INVENTION: Factors Involved in Gene Expression
; FILE REFERENCE: BB-1172
; CURRENT APPLICATION NUMBER: US/09/347,833
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,415
; EARLIER FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-347-833-4

Query Match 7.2%; Score 90; DB 3; Length 655;
Best Local Similarity 23.4%; Pred. No. 0.17;
Matches 47; Conservative 25; Mismatches 89; Indels 40; Gaps 8;
Qy 23 DRNAVTRCFA-----BEAFOAITDFNDLPSLFCACNVHOSVFEGESKEKFE 70
Db 280 DKEWYVCRAQKKSEREMELKEKEKNKEAADKQGTN-LYLKLNLDSDIDDEKLKEIFA 338
Qy 71 GLFRYDDCVTFQLFKSFRRYR-----INFSNPKSAARIELHETQFGKLLKLYFAOV 125
Db 339 D-FGTITSC---KVRDLNGVSGSGFVAFKSAEDASRALVAMNGKMTGSK--PLYVALA 392
Qy 126 QTPETDGDGLH:-----LAPPQAKQFLISPPSPVGVQWQIPNDATPVLYNLYAV 176
Db 393 QRKEERRARLQAQFSQMRPMVPPSVAPRPMYPPGVGVQQLFYGPQP------AF 445
Qy 177 AKLGPGEKYLHAGTSTSPV 197
Db 446 VNPQPGFGFQOHLIPGMRPSV 466

RESULT 12

US-09-562-737-12
; Sequence 12, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-12

Query Match 6.7%; Score 83.5; DB 4; Length 659;
Best Local Similarity 22.0%; Pred. No. 0.98;
Matches 54; Conservative 29; Mismatches 82; Indels 81; Gaps 12;
Qy 47 PNLFCACNVHOSVFEGESKEKFEGLFRYDDC-VTFOLEK--SFRRVRINFSNPKSAAR 103
Db 30 PNGRLTHDISLEHFEDELSE-----IIDCGISLQCLDTLSLRPRMGLLSAGSGN 82
Qy 104 ARIELHETQFGKLLKLYFAOVQTPETDGLKHLAPPQAKQFLISPPSPV----- 156
Db 83 AGSRL---QAEMQQMDLIDAAGRTPGAEDDESEDELAQAQRTGVGPPKAESVQDPAPRS 139
Qy 157 -GWQP-----INDATPV-----LNYDLLYAVAKLGPEKYE----- 186

us-10-030-613-1.rai

```

108 QY LHETQFRGKKLKYFAOYTPE-----TDGDKHLAPQPAPAKQLI 148
    |||      ||| |||      |||
43  DB LHE-----LYDLVTAPEDNEEASVQIFPDSVMLAQEGIDLLTFPPAGS---- 89
149 QY SPPSPPPVGWQIPINDATPVNLDLLAVAKLGKEKYLHA--CTESTPSVVHVHCD--- 203
    |||      ||| |||      |||
90  DB -PEEPHLSQP-----BQFORALGPVSMNLYPEVIDLTCH 125
204 QY -----SDIEEBDPKTSKPK-----LIQTRRGLPPSV 232
    ||| :|| :|| :||
126 DB EAGFPSPDDEEGVPVEPEPEPEPEPARTRPKMAPAI 167
    |||| :||

```

Search completed: August 22, 2003, 18:59:56
Job time : 32 secs

301 CTTCCAGCAATCTAAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTTCAG 360
361 AGGAAAAAATTAAGCTCTACTTTGACAGAGTTTCAGACTCCAGAGACAGATGGAGCAA 420
421 ACTGCACTTGGCTCCACCCGAGCTGCAAAACAGTTTCTCATCTCGCCGCCCTTCCCTCCC 480
481 ACCCTGTTGGCTGCGACGCCCATCAAGCATGCCACGCGAGTCTCAACTATGACCTCCCTTA 540
541 TGCTGTGGCCAACTAGGACAGGAGAGAGTATGAGCTCCATGCGAGGAGTGCAGTCCAC 600
601 CCCAAGTGTGCTGCTGACGCTGTCGACAGTGCACATAGAGGAGAGAGGACCCCAAGAC 660
661 TTCCCAAGCCAAATATCACTCAAACTCGGCGTCTGCGCTGCCACCTCCGCTGCCAA 720
721 CTGAGCTGCTGCTCTCTCGATAATAGCCGCTCTCTCTTTATCATGCTTTTCCGCCCT 780
781 GTTGTGTTGCAAAAAATTTGCTTTAAATTCCTGGGTGTTGGTGTGTTGAGATTCCTT 840
841 CCTGTTATCAAGCTCTCGCAAAAGGCTAGGAAAAAGTGCATATGCTCTCTGATCAT 900
901 ATCATACCAATTAAGTATACCCATTTTAGAA 934
1129 ATCATACCAATTAAGTATACCCATTTTAGAA 1162

RESULT 2
US-09-614-474-12
; Sequence 12, Application US/09614474
; Patent No. 6524819
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PC-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Inocyte ID No. 6524819 g1435039
US-09-614-474-12
Query Match 77.2%; Score 720.8; DB 4; Length 3184;
Best Local Similarity 95.9%; Pred. NO. 1.2e-228;
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

163 CTCGTTGTTTCGCTGCAATGTTCCACAGTCAGTCTTTTAAAGGAGAGAGAGAGGAAAA 222
222 CACTCTGTTGCCCTGTGTGGGATGTCGAGGCTTTTACCAATCAGGAGGTTAAGGAAAA 281
223 ATTTTCAGGAGGACTGTTTTCGAGCTTATGATGAGTGTGTGACGTTCCAGCTATTTAGAGTTT 282
282 ATTTTCGAGGACTGTTTTCGAGCTTATGATGAGTGTGTGACGTTCCAGCTATTTAGAGTTT 341
283 CAGAGCTGTCCGTATATAACTTTTCAGCAATCTTAATCTGACGCGGAGCTAGGATAGAGCT 342
342 CAGAGCTGTCCGTATATAACTTTTCAGCAATCTTAATCTGACGCGGAGCTAGGATAGAGCT 401
343 TCATGAACCCCAATTTTCAGAGGAGAAAAATTAAGCTCTTACTTTTTCACAGGTTTCAGACTCC 402
402 TCATGAACCCCAATTTTCAGAGGAGAAAAATTAAGCTCTTACTTTTTCACAGGTTTCAGACTCC 461
403 AGAGACAGATGAGAGCAAACTGACACTTGGCTCCACCCAGCTCCCAACAGTTTCTCAT 462
462 AGAGACAGATGAGAGCAAACTGACACTTGGCTCCACCCAGCTCCCAACAGTTTCTCAT 521
463 CTCGCCCCCTTCTCCACACCTGTTGGCTGGCAGGCCCATCAACGATGCCAGCCAGTCT 522
522 CTCGCCCCCTTCTCCACACCTGTTGGCTGGCAGGCCCATCAACGATGCCAGCCAGTCT 581
523 CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCCAGGAGAGAGTATGAGCTCCA 582
582 CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCCAGGAGAGTATGAGCTCCA 641
583 TGCGAGGACTGAGTCCACCCCAAGTGTCTGCTGCGACGCTGTGCGACAGTGCATAGAGGA 642
642 TGCGAGGACTGAGTCCACCCCAAGTGTCTGCTGCGACGCTGTGCGACAGTGCATAGAGGA 701
643 AGAAGAGGAGCCCAAGAGCTTTCCCAAGCAAAATATCATCCAACTCCGCGCTCCGCGCT 702
702 AGAAGAGGAGCCCAAGAGCTTTCCCAAGCAAAATATCATCCAACTCCGCGCTCCGCGCT 761
703 GCCACCTCGCTGCTCAACTGAGCTGCTGCTCTCGATTAATAGCGCTCTCTCTTT 762
762 GCCACCTCGCTGCTCAACTGAGCTGCTGCTCTCGATTAATAGCGCTCTCTCTTT 821
763 ATCATGCTTTTCCGCTGTTGTTTCAAAAAATTTCCCTTTAAATTCCTGGGTGTT 822
822 ATCATGCTTTTCCGCTGTTGTTTCAAAAAATTTCCCTTTAAATTCCTGGGTGTT 881
823 GGTGTTTTCAGATTTCTTCTTCTTATCAAGCTCTTCGACAAAGGCTAGGAAAAAGT 882
882 GGTGTTTTCAGATTTCTTCTTCTTATCAAGCTCTTCGACAAAGGCTAGGAAAAAGT 941
883 GATATGCTCTCTGATCATATCATACCCATTAAGTATACCCATTTTAGAA 934
942 GATATGCTCTCTGATCATATCATACCCATTAAGTATACCCATTTTAGAA 993

RESULT 3
US-09-614-474-3/c
; Sequence 3, Application US/09614474
; Patent No. 6524819
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PC-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Rattus norvegicus


```

; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: INCY
US-09-614-474-3

```

Query Match	45.2%	Score 422;	DB 4;	Length 1021;
Best Local Similarity	79.8%	Pred. No. 9.1e-130;		
Matches 624; Conservative	0;	Mismatches 140;	Indels	18; Gaps
QY	163	CTCGTTGTTGGCTGCAATGTTACCACTGTCAGTGTGTTGAAGGAGAGAGACAGGAAAA	222	
DB	823	CACCTCTGGTGCCTGTGTGGTGAATGTGGAGGTCTTTTACCAATCAGGAGGTTAAGGAAAA	764	
QY	223	ATTTGAGGAGCTGTTTCGGACTTTATGATGACTGTGTGACGTTTCCAGCTATTTAAGAGTTT	282	
DB	763	ATTTGAGGAGCTGTTTCGGAGCTATGACGAATGTGTGACGTTTTCAGCTGTTTAAGAGTTT	704	
QY	283	CAGAGCTG--TCGCTATAAATTTGACGAATCCTAAATTCGACCGC--AGCTAGATAGA	339	
DB	703	CGGACGGGTTCGGAATAAATTCAGCCACCCCAAGCTGCAGCCCGGTGCCCGATAGGA	644	
QY	340	GCCTCA--TGAACCCCAATTTCAGAGGGAATAAATTAAGCTCTACTTTGCACAGGTTTCAG	397	
DB	643	GCCTCATGTGGACCCAGTTCATGAGGGAAGAAGCTGAAACTCTACTTTGCACAGGTTTCAG	584	
QY	398	ACTCCAGAGACAGATGG-AGACAAACTGCA-CTTGGCTCCACCCAGCCCTGCCAAACAGT	455	
DB	583	ACCCACAGACAGATGCTAGACAACTGCAGTTTGGCACCCCAACAGCTTGCCTGCAACAGT	524	
QY	456	TTCTCATCTGCCCCCTTCTCTCCCACTGTTTGGCTGGCAGCCCACTCAACGATGCCACGC	515	
DB	523	TCTCATCTCACCCCTTTCATCT	574	
QY	516	CAGTCCCTCAAC-TATGACCTCTCTCTATGCTGTGGCCCAACTAGCACAGAGAGAGATAT	574	
DB	463	CAGTCCCTCAAGTACGACCTCTCTTATGCGTGGCCCAACTAGCACAGAGAGAGATATAT	404	
QY	575	GAGTCCCATCAGGAGCTGAGTCCACCCCAAGTGTGCTGTGACGCTGTGCGACAGTGCAC	634	
DB	403	GAGTGCATCGGGAACTGAGTCTACACGAGCGTGTGCTGTGACGCTGTGTGACAGCAG	344	
QY	635	ATAGAGGAAGAGGAGCCAAAGACTTCCCCAAAGCCAAATAATCATCCAGACCCGGCT	694	
DB	343	TTGGAGGAGGAGGAGATCCAAAGACTTCCCCAAAGCCAAATAATCATCCAGACCCGGCT	284	
QY	695	CCTGGCTGCCACCTCCGCTGTCGAAGTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	754	
DB	283	CCTGGCTGCCACCTCCGCTGTCGAAGTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	231	
QY	755	TCCTCTTTATCATGCTTTTTCCT	813	
DB	230	TCCTCT-TTATCATGCTTTTCT	172	
QY	814	TGGTGTTTGGTGTGTTGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	872	
DB	171	TGGGTATTTGGTGTGAGATTCCT	112	
QY	873	AGGAAAGGTGATGTCTCTCTGATCATATCATACCCATTAAAGTATATAACCCATTATTATAG	932	
DB	111	AGGAAAGGTGATGTCTCTCTGATCATATCATACCCATTAAAGTATATAACCCATTATTATAG	52	
QY	933	AA	934	
DB	51	AA	50	

RESULT 4
US-09-614-474-13
; Sequence 13, Application US/09614474
; Patent No. 6524819
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.

```

: APPLICANT: Edwards, Carla M.
: APPLICANT: Streeter, David G.
: TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
: FILE REFERENCE: PC-0013 US
: CURRENT APPLICATION NUMBER: US/09/614,474
: CURRENT FILING DATE: 2000-07-11
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PERL Program
: SEQ ID NO 13
: LENGTH: 828
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. 6524819 96017918
US-09-614-474-13

```

Query Match	29.9%;	Score 279.2;	DB 4;	Length 828;
Best Local Similarity	66.0%;	Pred. No. 2e-82;		
Matches 420;	Conservative 0;	Mismatches 213;	Indels 3;	Gaps 1;
QY	85	AGACAGGACGTGGGCTGTCACTCGTTGTTTTCGAGAAGAGCCTTTCAAGCAATCACTGA	144	
Db	85	AGAGAGGAAGAAGAGATGATTTTGGTGAATAATCAAGATGATTTCGATGAGATGATGGA	144	
QY	145	CTTCAATGACCTCCCAACTCGTTGTTTCGCTGCAATGTTCCACAGTCAGTGTGTGAAGG	204	
Db	145	TTTAAGTGATCTGCCTACCTCACCTTTTGCTTGCAGCGTCATGAAGCAGTGTGTGAGGC	204	
QY	205	AGAAGAGCAAGGAAAATTTAGGAGACTGTTTCGGACTTATGATGACTGTGTGACGTT	264	
Db	205	ACGAGAGCAGAAGAAAGATTTGAAGCACTCTTCACCATCTATGATGACCAGGTACTTT	264	
QY	265	CCAGCTATTTAAGAGTTTTCAGACGTGTCGCTATAAACTTCAGCAATCTTAATCTCGAGC	324	
Db	265	TCAGCTGTTTAAAAGCTTTAGAAGATCAGAATAAATTTTCAGCAAACTGAAAGCGCAGC	324	
QY	325	CCGAGCTAGGATGAGGCTTCATGAACCCCAATTCAGAGGGAAAAATTTAAAGCTCTACTTT	384	
Db	325	AAGAGCCGAATAGAACTCCACGAAACAGAGACTTCAATGGGCAGAACTAAAGCTATATTT	384	
QY	385	TGCACAGGTTTCAGACTCCAGAGACAGATGAGACAAACTGCATCTGGCTCCACCCAGCC	444	
Db	385	TGCACAGGTGCAGATGTCGCGCGAAGTGCGGGACAAGTCTATCTCTCGCGCCCGCAGCC	444	
QY	445	TGCCAAACAGTTTCTCATCTGCCCCCTTCTCTCCCACTGTGTGGCTGGCAGCCCCATCAA	504	
Db	445	TGTCAAGCAGTCTCTCATCTCCCTCAGCCTCTCCCCAGTGGGGTGGGAAGCAGAGCGA	504	
QY	505	CGATGCCACGCCAGTCTCTCAACTPATGACCTCTCTATGCTGTGGCCAAACTAGGACCCAG	564	
Db	505	AGATGCGATGCTGTATTAATPATGATTTACTCTGTGCTGTTTCCAAATTTGGGACCAAG	564	
QY	565	AGAGAATATGAGCTCCATGTCAGGGACTGAGTCCACCCCAAGTGTGCTGTCAGCTGTG	624	
Db	565	AGAGAATATGAACTTTCAGCGGGAAACAGATCGACACCCAGCGTGGTGTTCATGCTCTG	624	
QY	625	CGACAGTGCATAGAGGAAGAAGAGGACCCCAAGACTTCCCCAAAGGCCAAAATCATCCA	684	
Db	625	TGAAAGTGAACATGAGAGGAAGAAGAGAGCAAAAA---ACCCCAACAGAAAATTTGCCCA	681	
QY	685	AACTCGGCTCTCGGCTGCCACCTCCGTTGTCCAA	720	
Db	682	GACAAGGCCCGACCTCTCCGACCGCAGCGTTGAA	717	

RESULT 5
US-09-614-474-8
; Sequence 8, Application US/09614474
; Patent No. 6524819
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.

us-10-030-613-3.rni

Tue Aug 26 09:11:05 2003

```

; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PC-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6524819 206751_Rn.1
; US-09-614-474-8

Query Match      27.6%; Score 257.6; DB 4; Length 615;
Best Local Similarity 68.9%; Pred. No. 2.5e-75; Indels 0; Gaps 0;
Matches 353; Conservative 0; Mismatches 159;

QY 119 GAAGAAGCCCTTCAAGCAATCACTGACTTCAATGACCTCCCAACCTCGTTGTTGGGTGC 178
DB 103 GAAGACGATTGGAGAGATGATGACCTCAGCGACCTGCCACCTCCCTCTTGTGCTGC 162
QY 179 AATGTTCCAGTCAGTGTGTTGAAGAGAGAGAGCAAGAAATTTGAGGACGTGTT 238
DB 163 AGTGTCATGAAGCGGTGTTGAGGTTCCAGAGCAAGAGGAGGTTCCAGAGCCCTGTT 222
QY 239 CGGACTTATGATGACTGTGTGAGTTCAGCTATTTAAAGAGTTTCAGACGTGTCCTATA 298
DB 223 ACCCTCTATGATGACAGGTGCACATTCAGCTGTTCAAGAGCTTTCAGAGTGCAGATC 282
QY 299 AACTTCAGCAATCTAACTGACGAGCCGAGTATGAGTATGAGCTTCATGAACCCCAATTC 358
DB 283 AACTTCAGCAAGCTGAGGCTGCGGCGAGAGCAGGATCGAGTCCACGAGAGTGATTC 342
QY 359 AGAGGAAAAAATAAGCTTCTACTTTGCACAGCTTCAGACTCCAGAGACAGATGGAGAC 418
DB 343 CATGCCCGGAGCTGAAGCTTTACTTCGACAGGTGCAGGTTGTCGGGGAGGCTCGGGAC 402
QY 419 AAATGCACTTGGCTCCACCCAGCTGCGCAAGAGTTTCTCATCTGCCCTTCTCTCC 478
DB 403 AAGTCTTACTTACTGCCACACAGCCACCAAGAGTTCTCTATCTCCCTCCGCGCTCA 462
QY 479 CCACCTGTGCTGCGACCCCATCAAGATGCCAGCTCCCTCAACTATGACCTCCTC 538
DB 463 CCCCCCGTGGGTGGAAGCAGAGTGAGATGCCAGCGCAGTGATCAACTATGACCTCCTC 522
QY 539 TATGCTGTGGCCAACTAGGACCCAGGAGAGAGATGATGAGTCCATGAGGAGCTGATCC 598
DB 523 TGTGCGGTCTCCAAAGCTGGGGCCAGGGAGAAATACGAACCTGCACGCAACCGAGTCC 582
QY 599 ACCCAAGTGTGCTGTCGACGTGTCGACAG 630
DB 583 ACCCCAGTGTGCTGTCGACGTGTCGAGAG 614

RESULT 6
US-08-665-040-1
; Sequence 1, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUENTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
; TITLE OF INVENTION: FOR CHARACTERIZING IT.
; NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
ADDRESSEE: LADAS & PAREY
STREET: 26 WEST 61ST STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1 FOR DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,040
FILING DATE: JUNE 7, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: ES P9501140
FILING DATE: JUNE 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: JANET I. CORD
REGISTRATION NUMBER: 33,778
REFERENCE/DOCKET NUMBER: U010815-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
TELEFAX: (212) 246-8959
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA for mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: human
ORGANISM: Homo sapiens
DEVELOPMENTAL STAGE: foetal
TISSUE TYPE: Brain
IMMEDIATE SOURCE: gene library of cDNA
LIBRARY: gene library of cDNA from foetal
CLONE: BC-17.8-1 and BC-17.8-2
POSITION IN GENOME:
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9
MAP POSITION: 21q22.1-q22.2
FEATURE:
NAME/KEY: cDNA for mRNA, BC-17.8
LOCATION: 1..2174
FEATURE:
NAME/KEY: untranslated 5'
LOCATION: 1..48
FEATURE:
NAME/KEY: coding sequence
LOCATION: 49..560
OTHER INFORMATION: Down Syndrome critical
region 1 (DSCR1)
FEATURE:
NAME/KEY: DSCR1
LOCATION: 1..171 PEPTIDES
IDENTIFICATION METHOD: translation of the
OTHER INFORMATION: - deduced protein
OTHER INFORMATION: - proline-rich protein domains
OTHER INFORMATION: - glutamic acid-rich protein domains
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
FEATURE:
NAME/KEY: untranslated 3'
LOCATION: 564..2174
FEATURE:
NAME/KEY: 2 poly (A)
LOCATION: 1541..1546 AND 2132..2137
US-08-665-040-1
Query Match      19.0%; Score 177.4; DB 2; Length 2174;

```

1

us-10-030-613-3.rni

Tue Aug 26 09:11:05 2003

RESULT 10
US-08-635-066-1
Sequence 1, Application US/08635066
Patent No. 5945580
GENERAL INFORMATION:
APPLICANT: Dunsuir, Pamela
APPLICANT: Harpster, Mark H.
TITLE OF INVENTION: Capsicum Hemocellulase Polynucleotides
TITLE OF INVENTION: and Polypeptides
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/635,066
FILING DATE: 19-APR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 012176-005500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 609..2129
US-08-635-066-1

Query Match 4.4%; Score 40.8; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 0.0096;
Matches 18; Conservative 209; Mismatches 171; Indels 0; Gaps 0;
QY 393 TTCAGACTCCAGACAGATGGAGAACTGGCTCCACCCAGCCTGCCAAAC 452
Db 1031 TTCGAGCTGGCTGAGTGGAGTGGCTCCACCCAGCCTGCCAAAC 452
QY 453 AGTTCTCATCGCCCTTCTCCACCTGTGGCTGGCAGCCATCAACATGCCA 512
Db 1091 YY 1150
QY 513 CGCAGTCCTCAACTATGCTCTATGCTGGCCAACTAGGACGAGAGAGT 572
Db 1151 YY 1210
QY 573 ATGAGTCATGAGGAGTGGTCCACCCAGTGTCTGTCGTCGTCGTCGTCG 632
Db 1211 YY 1270
QY 633 ACATAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
Db 1271 YY 1330
QY 693 GTCCTGGCTGCCACCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 752
Db 1331 YY 1390
QY 753 TCTCTCTCTATCATGCTTTTCCCTGTTGTTGTC 790
Db 1391 YY 1428

US-08-232-463-14

RESULT 11
US-09-328-352-3446
Sequence 3446, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breston et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352

```

; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3446
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3446

```

	Query Match	3.5%	Score 33;	DB 4;	Length 1011;
	Best Local Similarity	52.6%;	Pred. No. 1.2;		
	Matches	72;	Conservative	0;	Mismatches 65; Indels 0; Gaps 0;
Qy	265	CCAGCTATTTAAGATTTCAGAGCTGCCGTATATAAACTTCAGCAATCCTTAATCTCGAGC	324		
Db	252	CCGCTCGCTTTTAAAGTAAAAACCGTGCTTTATCAGCTTTAAGCCAAAAATTAATTCATCA	311		
Qy	325	CCGAGCTAGGATAGAGCTTCATGAACCCCAANTTCAGAGGGAAAAATTAAGAGCTCTACTT	384		
Db	312	ACTGGCTATTAAAGAAATGATANTGCACAAACTACCGCCCAAAAAGTTATTCAATACAT	371		
Qy	385	TGCACAGGTTTCAGACTC	401		
Db	372	TGAACAAGATCATACTC	388		

```

RESULT 12
US-09-141-000-4/c
; Sequence 4, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING
; TITLE OF INVENTION: RECEPTOR PROTEINS
; FILE REFERENCE: 19999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-4

```

Query Match	3.5%;	Score 32.6;	DB 3;	Length 458;
Best Local Similarity	9.2%;	Pred. No. 0.97;		
Matches	39;	Conservative 125;	Mismatches 251;	Indels 0; Gaps 0;
Qy	228	AGGGACTGTTTCGGACTTATGACTGCTGTGAGCTTCCAGCTATTAAAGATTTTCAGAC	287	
Db	448	MDK.BMC..M.NR.D.BPMAS.Y...AK.KMCTYY.H.KD.CT.RH..T.D.BH..M.	389	
Qy	288	GTGTCGGTATAAAGCTTCAGCAATCTAAATCTGCAGCCGAGCTAGGATAGAGTTCATG	347	
Db	388	BT.BH..DKSHSNT.T.TM.AB...M..MKSMMRMB.....TNN.H..CT.MS.H.HK.RH	329	
Qy	348	AAACCAATTCAGAGGGAAAAATTAAGCTCTACTTTGCACAGGTTGCAGACTCCAGAGA	407	
Db	328	HTRB..H..SYRBBG....KWTs....SK.HT.S.AS.C..DMTWC..BB..YHT.HG	269	
Qy	408	CAGATGGAGACAAATGCACCTTGGCTCCACCCAGCCCTGCCAAACAGTTTCTCATCTCGC	467	
Db	268	AA.TM.HSH.BA.H.KTR.M...TB..DSB.MNR.MMT..B...N.R.SN.TH.YYMR	209	
Qy	468	CCCCTTCTCCCACTGTTGGCTGGCAGCCCATCAAGCATGCCACGCCAGCTCCTCAACT	527	
Db	208	.YCCYBYH.YBC.M.KCBM.GMK.YGT.GSMYYMA...G.NAT.GSR..NC..AYM..TM	149	
Qy	528	ATGACCTCCCTATGCTGTGGCCAAACTAGGACCCAGGAGAGTATGAGTCCATGCAG	587	
Db	148	GT..GSTBCRDRCST.HCSB.G.YM.KSN.KR.GMA..H.B.ASSGHR.MYB.CSCC...	89	
Qy	588	GGACTGAGTCCACCCCAAGTGTGCTGTCGACGTTGTCGACAGTGCATAGAGG	641	

[illegible]

```

? LENGTH: 15652
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)_(15652)
? OTHER INFORMATION: n = A, T, G, or C
? NAME/KEY: misc_feature
? LOCATION: (1)...(15652)
? OTHER INFORMATION: r = G or A; y = T/U or C; m = A or C;
? OTHER INFORMATION: k = G or T/U; s = G or C; w = A or T/U;
? OTHER INFORMATION: b = G, C, or T/U; d = A, G, or T/U;
? OTHER INFORMATION: h = A, C, or T/U; v = A, G, or C
? NAME/KEY: exon
? LOCATION: (3419)...(3444)
? NAME/KEY: intron
? LOCATION: (3445)...(3908)
? NAME/KEY: exon
? LOCATION: (3909)...(3993)
? NAME/KEY: intron
? LOCATION: (3994)...(4694)
? NAME/KEY: exon
? LOCATION: (4695)...(4898)
? NAME/KEY: intron
? LOCATION: (4899)...(5652)
? NAME/KEY: exon
? LOCATION: (5653)...(5838)
? NAME/KEY: intron
? LOCATION: (5839)...(7184)
? NAME/KEY: exon
? LOCATION: (7185)...(7205)
? NAME/KEY: intron
? LOCATION: (7206)...(8310)
? NAME/KEY: exon
? LOCATION: (8311)...(8806)
? NAME/KEY: intron
? LOCATION: (8807)...(12271)
? NAME/KEY: exon
? LOCATION: (12272)...(12406)
? NAME/KEY: intron
? LOCATION: (12407)...(12820)
? NAME/KEY: exon
? LOCATION: (12821)...(12991)
? NAME/KEY: intron
? LOCATION: (12992)...(14089)
? NAME/KEY: exon

```

us-10-030-613-3.rni

Tue Aug 26 09:11:05 2003

LOCATION: (14090)...(14191)
 NAME/KEY: intron
 LOCATION: (14192)...(14477)
 NAME/KEY: exon
 LOCATION: (14478)...(14543)
 NAME/KEY: intron
 LOCATION: (14544)...(15002)
 NAME/KEY: exon
 LOCATION: (15003)...(15194)
 NAME/KEY: intron
 LOCATION: (15195)...(15652)
 US-09-422-936-60

Query Match 3.4%; Score 31.6; DB 4; Length 15652;
 Best Local Similarity 53.2%; Pred. No. 17; Indels 0; Gaps 0;
 Matches 67; Conservative 0; Mismatches 59;
 QY 648 AGGACCAAGACTTCCCAAGCCCAAAATCATCCAAACTCGGGTCTGCGCTGCCAC 707
 Db 9629 AAGTCCATAGGTGCCCCAGATCTCTCATTTATCCACGCTACCTCTGCCCTCTC 9570
 QY 708 CCTCGGTGCAACTGAGTGGCTCTCTCGATAATAGCGCTCTCTTTATCAT 767
 Db 9569 CCCACTGCCATCTTTTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9510
 QY 768 GCTTTT 773
 Db 9509 GCTATT 9504

RESULT 14
 US-09-619-353-13
 Sequence 13, Application US/09619353
 Patent No. 6410249
 GENERAL INFORMATION:
 APPLICANT: Ngai, John
 APPLICANT: Specs, David J.
 APPLICANT: Lin, David M.
 APPLICANT: Isacoff, Ehud Y.
 APPLICANT: Dittman, Andrew H.
 APPLICANT: Fan, Jinhong
 TITLE OF INVENTION: Odorant Receptors
 FILE REFERENCE: B99-038-2
 CURRENT APPLICATION NUMBER: US/09/619,353
 CURRENT FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 60/144,766
 PRIOR FILING DATE: 1999-07-20
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 13
 LENGTH: 2595
 TYPE: DNA
 ORGANISM: Brachydanio rerio (zebrafish)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (4)...(2592)
 US-09-619-353-13

Query Match 3.4%; Score 31.4; DB 4; Length 2595;
 Best Local Similarity 54.9%; Pred. No. 6.9;
 Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 353 CAATTGAGAGGAAATTAAGCTCTACTTTCACAGGTTCAGACTCCAGAGACAGAT 412
 Db 648 CATTACAGATGAGATTATGAGCTTCTGCTTTCAGAGTTTCATACAGCACACCGAAC 707
 QY 413 GGAGACACTCTGCTTGGCTCCAGCCAGCTCCCAACAGTTTCTCATCTC 465
 Db 708 GGAGGGAATTTGATCGCTTTAAGCAATCTCTCCAGACTCTACTAGCAGATC 760

RESULT 15
 US-09-791-211-10

Sequence 10, Application US/09791211
 Patent No. 6448080
 GENERAL INFORMATION:
 APPLICANT: Donna T. Ward
 APPLICANT: Andrew T. Watt
 TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
 FILE REFERENCE: RTS-0205
 CURRENT APPLICATION NUMBER: US/09/791,211
 CURRENT FILING DATE: 2001-02-23
 NUMBER OF SEQ ID NOS: 90
 SEQ ID NO 10
 LENGTH: 9884
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: unsure
 LOCATION: 24962
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 64383
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 65468
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 65469
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 65470
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 65471
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 87130
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 89049
 OTHER INFORMATION: unknown
 OTHER INFORMATION:
 US-09-791-211-10

Query Match 3.4%; Score 31.4; DB 4; Length 9884;
 Best Local Similarity 45.1%; Pred. No. 60;
 Matches 116; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
 QY 121 AGAAGCTTTCAAGCAATCACTGACTTCAATGACCTCCCACTGTTGTTGGTGCAA 180
 Db 79522 AGACCCCATCTCTGAAAAAATGAGTGCATGGTACCTTTACAGTGTAGTTA 79581
 QY 181 TGTTCCAGCTCAGTGTGTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 Db 79582 TGAAG 79641
 QY 241 GACTTATGATGACTGTGTGACGTTTCAGCTATTTAAGAGTTTCAGACGTTCCGTATA 300
 Db 79642 GAGTGGTGGCTCACACCTGTAATCCAGCAGCTCTGGAGGCGGCGGCTGATCATGA 79701
 QY 301 CTTGAGCAATCCCTAAATCTCGAGCCGAGCTAGGATAGAGCTTTCATGAAACCCCAATTCAG 360
 Db 79702 GGTCCAGGAGATCGAGACCATCTCTGGGAGGCGGCGGCTGATCATGA 79761
 QY 361 AGGGAAAAAATTAAGC 377
 Db 79762 AAAAGAAAAAATTAAGC 79778

Search completed: August 24, 2003, 12:13:07
 Job time : 100 secs

QY 181 PGKYLHAGTSTPSVHVHVCDSIEEDPKTSKPKIIQTRRPGLPSPVSN 234
 DB 202 PGKYLHAGTSTPSVHVHVCDSIEEDPKTSKPKIIQTRRPGLPSPVSN 255
 RESULT 3
 AAE18916
 ID AAE18916 standard; Protein; 192 AA.
 XX AAE18916;
 XX 17-MAY-2002 (first entry)
 XX Human MCIP associated protein #3.
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human.
 KW Homo sapiens.
 OS WO200204491-A2.
 PN 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-US21662.
 PF 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX Williams SR, Rothermel B;
 WPI: 2002-179698/23.
 DR N-PSDB; AAD30157.
 XX Screening for modulators of muscle calcineurin interacting protein
 (MCIP) binding, expression or phosphorylation, useful for treating
 cardiac hypertrophy or heart failure, comprises mixing MCIP,
 calcineurin and a test compound -
 Example 1; Page 165-166; 174pp; English.
 The invention relates to muscle calcineurin interacting proteins (MCIPs)
 and nucleic acid molecules encoding such proteins. MCIPs form a physical
 complex with the catalytic subunit of calcineurin and increased levels
 of MCIPs correspond to a reduced ability of calcineurin to stimulate
 transcription of certain target genes. The invention also relates to
 methods for identifying modulators of MCIP binding, expression or
 phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 may be used for treating cardiac hypertrophy and heart failure.
 Antibodies to MCIP can be used in characterizing the MCIP content of
 healthy and diseased tissues and subsequently for determining the
 presence or absence of cardiomyopathy or as predictor of heart disease.
 The present sequence is human MCIP associated protein.
 Note: This sequence has been described as human MCIP splice variant in
 specification, however the sequence seems to be a MCIP associated
 protein.
 XX Sequence 192 AA;
 SQ
 Query Match 72.8%; Score 913; DB 23; Length 192;
 Best Local Similarity 93.0%; Pred. No. 3.3e-89;
 Matches 174; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 48 NSLFACNVHQSVEGESKKEKFLFRDYDCVTFQLEKSFRRVNFNSKSAARARIE 107
 DB 6 STLACVVDVEFTNQVEKKEKFLFRDYDCVTFQLEKSFRRVNFNSKSAARARIE 65
 QY 108 LHETQFRGKKLYFAQVOTPETDGDKHLAPQAKQFLISPPSSPPVSNQINDATPV 167

Db 66 LHETQFRGKKLYFAQVOTPETDGDKHLAPQAKQFLISPPSSPPVSNQINDATPV 125
 QY 168 LNYDLLYAVAKLPGKEKYLHAGTSTPSVHVHVCDSIEEDPKTSKPKIIQTRRPG 227
 Db 126 LNYDLLYAVAKLPGKEKYLHAGTSTPSVHVHVCDSIEEDPKTSKPKIIQTRRPG 185
 QY 228 LPPSVSN 234
 Db 186 LPPSVSN 192
 RESULT 4
 AAE18913
 ID AAE18913 standard; Protein; 197 AA.
 XX AAE18913;
 XX 17-MAY-2002 (first entry)
 XX Mouse MCIP associated protein #4.
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; mouse.
 KW Mus musculus.
 OS WO200204491-A2.
 PN 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-US21662.
 PF 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX Williams SR, Rothermel B;
 WPI: 2002-179698/23.
 DR N-PSDB; AAD30154.
 XX Screening for modulators of muscle calcineurin interacting protein
 (MCIP) binding, expression or phosphorylation, useful for treating
 cardiac hypertrophy or heart failure, comprises mixing MCIP,
 calcineurin and a test compound -
 Disclosure; Page 154; 174pp; English.
 The invention relates to muscle calcineurin interacting proteins (MCIPs)
 and nucleic acid molecules encoding such proteins. MCIPs form a physical
 complex with the catalytic subunit of calcineurin and increased levels
 of MCIPs correspond to a reduced ability of calcineurin to stimulate
 transcription of certain target genes. The invention also relates to
 methods for identifying modulators of MCIP binding, expression or
 phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 may be used for treating cardiac hypertrophy and heart failure.
 Antibodies to MCIP can be used in characterizing the MCIP content of
 healthy and diseased tissues and subsequently for determining the
 presence or absence of cardiomyopathy or as predictor of heart disease.
 The present sequence is mouse MCIP associated protein.
 Note: This sequence has been described as mouse MCIP2 encoding DNA in
 the specification, however the sequence seems to be a MCIP associated
 protein.
 XX Sequence 197 AA;
 SQ
 Query Match 72.2%; Score 905; DB 23; Length 197;
 Best Local Similarity 91.4%; Pred. No. 2.4e-88;
 Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

[illegible]

RESULT 6:	
AAE18918	
ID	AAE18918 standard; Protein; 212 AA.
XX	
XX	
AC	AAE18918;
XX	
XX	
DT	17-MAY-2002 (first entry)
XX	
DE	Human MCIP associated protein #5.
XX	
XX	
XX	
KW	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy; heart disease; human.

Homo sapiens.
 WO200204491-A2.
 17-JAN-2002.
 06-JUL-2001; 2001WO-US21662.
 07-JUL-2000; 2000US-216601P.
 13-FEB-2001; 2001US-0782953.
 (TEXA) UNIV TEXAS SYSTEM.
 (WILL/) WILLIAMS S R.
 (ROTH/) ROTHERMEL B.
 Williams SR, Rothermel B;
 WPI: 2002-179698/23.
 N-PSDB: AAD30159.
 Screening for modulators of muscle calcineurin interacting protein
 (MCIP) binding, expression or phosphorylation, useful for treating
 cardiac hypertrophy or heart failure, comprises mixing MCIP,
 calcineurin and a test compound -
 Disclosure: Page 171-172; 174pp; English.
 The invention relates to muscle calcineurin interacting proteins (MCIPs)
 and nucleic acid molecules encoding such proteins. MCIPs form a physical
 complex with the catalytic subunit of calcineurin and increased levels
 of MCIPs correspond to a reduced ability of calcineurin to stimulate
 transcription of certain target genes. The invention also relates to
 methods for identifying modulators of MCIP binding, expression or
 phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 may be used for treating cardiac hypertrophy and heart failure.
 Antibodies to MCIP can be used in characterising the MCIP content of

QY	48	NSLFACNVHQSVPFEGESKEFGLFRTYDDCVTFQLFKSFRRVRINFNSNPKSAARARIE	107
Db	11	STLVACVVDEVTNQEVKEFGLFRTYDECVTQLFKSFRRVRINFNSHFKSAARARIE	70
QY	108	LHETQFRGKKLKYFAQVQTPETDGDGLHLPQPAKQFLISPSSPPVGVQPIINDATPV	167
Db	71	LHETQFRGKKLKYFAQVQTPETDGDGLHLPQPAKQFLISPSSPPVGVWKPISDATPV	130
QY	168	LNVDLLVAVAKLGPGEKVELHAGTESTPSVVHVCDSDIEEEDDKTSPKPIIOTRRPG	227
Db	131	LNVDLLVAVAKLGPGEKVELHAGTESTPSVVHVCDSDMEEBEDPKTSPKPIIOTRRPG	190
QY	228	LPSPSVN	234
Db	191	LPSPSVN	197

RESULT 5
AAE18917

AC	AAE18917;
XX	
XX	17-MAY-2002 (first entry)
DT	
XX	Human MCIP associated protein #4.
DE	
XX	
XX	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW	heart failure; cardiomyopathy; heart disease; human.
KW	
XX	
XX	Homo sapiens.
OS	
XX	

17-JAN-2002.

06-JUL-2001; 2001WO-US211662.

07-JUL-2000; 2000US-216601P.

13-FEB-2001; 2001US-0782953.

(TEXA) UNIV TEXAS SYSTEM.
(WILL/) WILLIAMS S R.
(ROTH/) ROTHERMEL B.

Williams SR, Rothermel B;
WPI; 2002-179698/23.
N-PSDB; AAD30158.

Screening for modulators of muscle calcineurin interacting protein (MCIP) binding, expression or phosphorylation, useful for treating cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -

Example 1; Page 168-169; 174pp: English.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterizing the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence has been described as splice variant of MCIP1 Note: this sequence has been described as splice variant of MCIP1 initiated by exon 4 in the specification, however the sequence seems to be a MCIP associated protein.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to

XX Williams SR, Rothermel B;
PI
XX
DR WPI; 2002-179698/23.
DR N-PSDB; AAD30152.
XX
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
PS Disclosure; Page 148-149; 174pp; English.
XX
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC complex with the catalytic subunit of calcineurin and increased levels
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
CC transcription of certain target genes. The invention also relates to
CC methods for identifying modulators of MCIP binding, expression or
CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or
CC may be used for treating cardiac hypertrophy and heart failure.
CC Antibodies to MCIP can be used in characterising the MCIP content of
CC healthy and diseased tissues and subsequently for determining the
CC presence or absence of cardiomyopathy or as predictor of heart disease.
CC The present sequence is human MCIP associated protein.
XX
SO

[illegible]

RESULT 11	
AAW73898	
ID AAW73898	standard; Protein; 171 AA.
XX	
AC AAW73898;	
XX	
DT 09-APR-1999	(first entry)
XX	
DE Human Down's Syndrome	critical region 1 protein.
XX	
KW .DSCR1; Down's Syndrome	critical region 1; human; pathogenic abnormality;
KW Central Nervous System	development; mental retardation; heart defect.
OS Homo sapiens.	
XX	
PN US5869318-A.	
XX	
PD 09-FEB-1999.	
XX	
PF 07-JUN-1996;	96US-0665040.
XX	
PR 07-JUN-1995;	95ES-0001140.
XX	
PA (PALL/) PALLEJA X E.	

XX	
PI	Fuentes JU, Pallega XE, Pritchard M;
DR	WPI; 1999-152781/13.
DR	N-FSDB; AAX01282.
XX	
PT	DNA encoding foetal brain proteins - believed to be associated with
PT	Down's syndrome
XX	
PS	Claim 4; Column 15-18; 19pp; English.
XX	
CC	This sequence is encoded by the Down's Syndrome critical region 1 (DSCR1)
CC	gene of the invention. The DSCR1 gene was found to be located in the
CC	q22.1-22.2 region of human chromosome 21. An increase in the
CC	expression of DSCR1 mRNA in the brains of young rats, compared to
CC	expression levels in the brains of adult rats, compared to
CC	for DSCR1 during the development of adult rats, suggests an important role
CC	that over expression of DSCR1 may be involved in pathogenic abnormalities
CC	of mental retardation and/or heart defects as found in Down's syndrome
CC	patients.
XX	
SQ	Sequence 171 AA:
XX	

```

Query Match          46.3%:   Score 581;   DB 20;   Length 171;
Best Local Similarity 67.1%:   Pred. No. 9.3e-54;
Matches 112; Conservative 22; Mismatches 25; Indels      8;

Qy    68 KFEGLFRTYDDCVTFOLFKFRRVRINFSNPKSAARIELHETQFRGKKLKLY
Db    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    5 KFESLFRTYDRDITFYFKSFKRVIRNFSPNSAADRLQHKTEFLGKEMKLY
Db    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    128 PETYDGDKHLAPQAPQAFILSPSSPPVGVQWQINDATPVNLVDLLAYAVAKLGP
Db    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    65 IGSS-----HLAPNPQKFLISPSAPPVGYKVEDATPVINYDLLYAISKLG
Db    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    188 HACTESTPVSVVHVCDSDIEBEDPKTS----PKPIIQTRRPGLP 230
Db    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 HAATDTTPSVVVHVCSDEQEKEEEMERMRPKPIQTRRPETP 167

RESULT 12
ABB71467
ID ABB71467 standard; Protein; 292 AA.
XX
```

RESULT 12

ABB71467	
ID	ABB71467 standard; Protein; 292 AA.
XX	
AC	ABB71467;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 41193.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
FA	(PEKE) PE CORP NY.
XX	
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
DR	N-PSDB; ABL15570.
XX	
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or
PT	genes from Drosophila and for elucidating cell signalling and cell

us-10-030-613-1.rag

Tue Aug 26 09:11:01 2003

PS Claim 13; SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number

CC of 5' ESTs derived from total human RNAs or polyA+ RNAs derived from 30

CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'

CC untranslated region (UTR) of the mRNA because they are not well suited for

CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for

CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in

CC those cases where longer cDNA sequences have been obtained, the full 5'

CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'

CC ends and can therefore be used to obtain full length cDNAs and genomic

CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and

CC chromosome mapping procedures. They are used to obtain upstream

CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 111 AA;

SQ Query Match 25.4%; Score 318; DB 21; Length 111;

Best Local Similarity 57.7%; Pred. No. 6.9e-26;

Matches 64; Conservative 16; Mismatches 27; Indels 4; Gaps 1;

QY 42 DFNDLPNSLFACNVHQSVFEGESKEKFGFLRTYDCVTFQLFKSFRRVRINFSNPKSA 101

DB 5 NFNYFSFLIACVANSIDFSESETRAKFESLFRTYDKDITFQYFKSKVRINFSNPKSA 64

QY 102 ARARTELHETOFGRKKLKLKLYFAQVOTPETDGLKHLAPPQPAKOFLLISPPS 152

DB 65 ADARLQLHKTEPLGKEMKLYFAQYLHGSS----HLAPPNPKQFLISPPA 111

RESULT 14

AAB58452

ID AAB58452 standard; Protein; 142 AA.

XX AC AAB58452;

XX 14-MAR-2001 (first entry)

XX Lung cancer associated polypeptide sequence SEQ ID 790.

DE Human; lung cancer associated protein; neuroprotective; cytostatic;

XX cardioactive; immunomodulatory; muscular active; vulnerary;

KW gastrointestinal; nephrotropic; antiinfective; gynecological;

KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

KW proliferative disorder; wound healing; infectious disease.

XX Homo sapiens.

OS WO2000055180-A2.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

XX N-PSDB; AAF18328.

XX Lung cancer associated gene sequences, referred to as lung cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

PT such as lung cancer -

XX Claim 11; Page 1325-1326; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer

PT Interactions -

XX Disclosure; SEQ ID NO 41193; 21pp + Sequence Listing; English.

PS The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 292 AA;

SQ Query Match 34.5%; Score 432.5; DB 22; Length 292;

Best Local Similarity 41.1%; Pred. No. 1.5e-37;

Matches 92; Conservative 36; Mismatches 75; Indels 21; Gaps 6;

QY 4 PQQGHVPEDGGLFLLCCIDRDWATRCFAEFAQAITDFNDLPNSLFACNVHQSVFEGE 63

DB 86 PNOHPSLPKEGD-----VDS-----TEPEVDA-DSFDLPSTSIIVTNIHSEVFANP 131

QY 64 ESKEKFGFLRTYDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFGRKK-LKLYF 122

DB 132 ELKAMEELFRIFSESATFQWLSFRRLRVNYDNAIAANARIKLHQYEENKKTIVTCYF 191

QY 123 AQVOTPETDGLKHLAPPQPAKOFLLISPPSPVGPQINDATPVNLDLLYAVAKLGP 182

DB 192 AQVPTPVSN---KNLQPPAPVKQFLISPPSPAGWEPREGEPLVNHDLAALSLTPG 248

QY 183 EKVHELHAGTSTPSVVVHVCDSIDIEEDPKTSKPKKTIOTRRP 226

DB 249 ESHELHQSEDPALIVHTA---MLAETGGLQVKAPIVQTKCP 289

RESULT 13

AAG01768

ID AAG01768 standard; Protein; 111 AA.

XX AC AAG01768;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5849.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping.

XX Homo sapiens.

OS EP1033401-A2.

PN 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC01774.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX

human genome-derived single exon nucleic acid probes useful for analysing gene expression in human adult liver -

Search completed: August 22, 2003, 18:56:09
Job time : 87 secs

A:Accession: T49641
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-315 <SCH>
 A:Cross-references: EMBL:AL355932; GSPDB:GN00116; NCSP:B5022.230
 A:Experimental source: BAC clone B5022; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B5022.230
 A:Map position: 6

Query Match 11.6%; Score 146; DB 2; Length 315;
 Best Local Similarity 31.0%; Pred. No. 2e-05;
 Matches 49; Conservative 21; Mismatches 60; Indels 28; Gaps 7;

QY 42 DFNDLP-----NSLFACNVHQ-SVFEGESKEFGLFRYDDCV--TFOLF---- 85
 DB DUSNUPPQSPPPPSNTLFTNNSLDVFSADN-----LQIRDLISQTAFIAWSP 99
 QY 86 -KSFRVRINFNPKSAARARIELHETQFRGKKLYFAQVQTPETDGDKLHLAPPOPAK 144
 DB LKSFRIIVTFDEQAATAVRSVMDGEAILGRCRVYFQPTPIDVSAADKHLALPDAGK 159
 QY 145 QFLISPPSPVGVW-OPINDA--TPVLNVDLLYAVAKL 179
 DB LFFISPPSPPHDWSQRMEDAPNTVMHAEADLAELAKL 197

RESULT 3

S62444 conserved hypothetical protein SPAC13G6.15c - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: hypothetical protein SPAC24B11.04c
 C:Species: Schizosaccharomyces pombe
 C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 10-Dec-1999
 A:Accession: S62444; S62549; T37650; T38331
 R:Odell, C.; Bowman, S.
 Submitted to the EMBL Data Library, October 1995
 A:Reference number: S62430
 A:Accession: S62444
 A:Molecule type: DNA
 A:Residues: 1-163 <ODE>
 A:Cross-references: EMBL:Z54308; NID:g1008985; PIDN:CAA91108.1; PID:g1009000
 R:Odell, C.; Churcher, C.M.
 Submitted to the EMBL Data Library, November 1995
 A:Reference number: S62546
 A:Accession: S62549
 A:Molecule type: DNA
 A:Residues: 1-163 <OD2>
 A:Cross-references: EMBL:Z67757; NID:g1061288; PIDN:CAA91769.1; PID:g1061292
 R:Odell, C.; Bowman, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.
 Submitted to the EMBL Data Library, October 1995
 A:Reference number: Z21734
 A:Accession: T37650
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-163 <OD3>
 A:Cross-references: EMBL:Z54308; PIDN:CAA91108.1; GSPDB:GN000066; SPDB:SPAC13G6.15c
 R:Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 Submitted to the EMBL Data Library, November 1995
 A:Reference number: Z21786
 A:Accession: T38331
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-163 <OD4>
 A:Cross-references: EMBL:Z67757; PIDN:CAA91769.1; GSPDB:GN000066; SPDB:SPAC24B11.04c
 A:Experimental source: strain 972h-; cosmid c24B11
 C:Genetics:
 A:Gene: SPAC24B11.04c; SPAC13G6.15c; SPAC24B11.04c
 A:Map position: 1L
 C:Superfamily: fission yeast hypothetical protein SPAC13G6.15c

Query Match 8.6%; Score 107.5; DB 2; Length 163;
 Best Local Similarity 34.9%; Pred. No. 0.024;

Matches 29; Conservative 13; Mismatches 28; Indels 13; Gaps 3;
 QY 137 LAPPOPAKOFLLISPPSPVGVQPIINDATP---VLNVDLLYAVAKLGPGEKVELHAGTES 193
 DB LQVPKFEKNWLLISPPSPVGVQPIINDATP---VLNVDLLYAVAKLGPGEKVELHAGTES 193
 QY 194 TPSVVVHVCDSDIEEEDPKTSP 216
 DB LQVPKFEKNWLLISPPSPVGVQPIINDATP---VLNVDLLYAVAKLGPGEKVELHAGTES 193

RESULT 4

A56954 yes-associated protein, 65K - human

N:Alternate names: yap65 protein
 C:Species: Homo sapiens (man)
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 23-Sep-2002
 C:Accession: A56954; S46973
 R:Sudol, M.; Bork, P.; Einbond, A.; Kastury, K.; Druck, T.; Negrini, M.; Huebner, K.; J. Biol. Chem. 270, 14733-14741, 1995
 A:Title: Characterization of the mammalian YAP (Yes-associated protein) gene and its A:Reference number: A56954; MUID:95301570; PMID:7782338
 A:Accession: A56954
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-454 <SUD>
 A:Cross-references: EMBL:X80507; NID:g517176; PIDN:CAA56672.1; PID:g517177
 C:Genetics:
 A:Gene: YAP
 C:Superfamily: yes-associated protein; WW repeat homology
 F:1/1-208/Domain: WW repeat homology <WW1>

Query Match 7.8%; Score 98; DB 2; Length 454;
 Best Local Similarity 31.1%; Pred. No. 0.62;
 Matches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;

QY 139 PPQPAKQFLISPPSPVGVQPIINDATPVLNVDLLYAVAKLGPGEKVELHAGTESPSV- 197
 DB PPQPAKQFLISPPSPVGVQPIINDATPVLNVDLLYAVAKLGPGEKVELHAGTESPSV- 197
 QY 198 -----VVHV-CDSDIEEEE-----DPKSPKPKIITRRRGLPPS 231
 DB PPAQVQVGVQPIINDATPVLNVDLLYAVAKLGPGEKVELHAGTESPSV- 197

RESULT 5

T37629 hypothetical protein SPAC13F5.04c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37629
 R:Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 Submitted to the EMBL Data Library, September 1997
 A:Reference number: Z21733
 A:Accession: T37629
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-277 <BRO>
 A:Cross-references: EMBL:Z99091; PIDN:CAB11767.1; GSPDB:GN000066; SPDB:SPAC13F5.04c
 A:Experimental source: strain 972h-; cosmid c13F5
 C:Genetics:
 A:Gene: SPDB:SPAC13F5.04c
 A:Map position: 1
 A:Introns: 20/3

Query Match 7.7%; Score 96; DB 2; Length 277;
 Best Local Similarity 22.4%; Pred. No. 0.5;
 Matches 46; Conservative 22; Mismatches 77; Indels 60; Gaps 7;

QY 79 CVTFQFLKSFRRVRINF-----SNPKSAAR-----ARIELHETOPRGK 116
 DB CFFPRVSSFAILRISFLSIKFKCKLRALQVANPQKAKTPSNHATELOQSSNST 61

QY 117 KLKLYFAOVQTPETDGLKHLAPP--QPAKQFLISPPSPVGVQWP--INDA 164
 Db 62 TLPTQEAARVETNASASHETSFALPTTSPAASLSISPTKSAVSEPNVADYKSLSTPA 121
 QY 165 TPVNLNLYLAVAKLGPGEYK-----LHAGTSTPSV-----VHVCDSDIEE 208
 Db 122 APQLN-----SPSHSYETPTPTSTITENLPTIDTRSTRSSHIQSLSPESKQ 171
 QY 209 EEDPKTSPKPKIQTTRRPGLPSPVS 233
 Db 172 TSDGHRPPSTSTITSTSIDPSVA 196

RESULT 6
 A72203
 cellobiose-phosphorylase - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: A72203
 R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: A72203
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-813 <ARN>
 A:Cross-references: GB:AE001822; GB:AE000512; NID:g4982429; PIDN:AND36910.1; PID:g498243
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1848

Query Match 7.5%; Score 94; DB 2; Length 813;
 Best Local Similarity 24.5%; Pred. No. 3;
 Matches 45; Conservative 27; Mismatches 78; Indels 34; Gaps 9;

QY 51 FACNVHOSVFEGEESKEFGLFRPYDCVTFQL--FKSFRVRVNFNSPKSAARARIEL 9;
 Db 126 FTGEVHYLVLENKAEPKRIKLFSEFELNWDNDTNRQ---NYSTGEVIEGSIY 108
 QY 109 HETQTRGKKLYFAOVQTP-----ETDGDK-LHAPPQPAKQFLI--SPSPSPVGVQWPI 182
 Db 183 HKTEYRRNRHYAFVSNQPIDGPDTRDFESFGLYSGFEAPQAVVEGKPRNSVASGAPI 161
 QY 162 NDATPVNLNLYLAVAKLGPGEYKELHAGTSTPSVVHVCDSDIEEEDPKTSPKPKII 242
 Db 243 ASH-----YLEIELAPSEKKEL-----IFIL-CYVENPEERKE-KPGVI 221
 QY 222 QTRR 225
 Db 281 NKRR 284

RESULT 7
 AC2098
 hypothetical protein alr2338 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. strain PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AC2098
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 C:Accession: AC2098
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-751 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA874037.1; PID:g17131430; GSPDB:GN00179
 A:Experimental source: strain PCC 7120

C:Genetics:
 A:Gene: alr2338

Query Match 7.5%; Score 93.5; DB 2; Length 751;
 Best Local Similarity 22.0%; Pred. No. 3;
 Matches 48; Conservative 30; Mismatches 95; Indels 45; Gaps 7;

QY 44 NDLPNSLFACNVHOSVFEGEESKEFGLFRPYDCVTFQLFKSFRVRVNFNSPKSAAR 103
 Db 109 NOTESQIAALKSAQSIGPAIENNEQVEKMLGSLDNLTI-AFNSLQQLKKNYQNSLSQOL 167
 QY 104 ARIELHETO-----FRGKKLYFAOVQTPETDGLKHLAPPQPAKQFLISPPS 152
 Db 168 SRMYNQOQGETIVEELIDRLUGELT-----AIQETSTAKAQLSPPTVLOPPELOPPS 221
 QY 153 SPV-----GWOPINDATPVNLNLYLAVAKLGPGEYKELHAGTSTPSV 197
 Db 222 SPVVNLSPPTVLOFPDQSQNPLOASTPLEETSTTKPSVSITPPEK-----STPVT 273
 QY 198 VHVCDSDIEEED--PKTSP--KPKIQTTRRPGLPSPS 231
 Db 274 IVPPPQETRPETKSVIPKVSFDPSETKLOSSEKAAEAPS 311

RESULT 8
 T19033
 hypothetical protein C07A4.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T19033
 R:White, S.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z19063
 A:Accession: T19033
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-270 <WIL>
 A:Cross-references: EMBL:Z78536; PIDN:CAB01717.1; GSPDB:GN00028; CESP:C07A4.1
 A:Experimental source: clone C07A4
 C:Genetics:
 A:Gene: CESP:C07A4.1
 A:Map position: X
 A:Introns: 54/2; 79/3; 119/3; 142/1

Query Match 7.3%; Score 92; DB 2; Length 270;
 Best Local Similarity 24.6%; Pred. No. 1.1;
 Matches 42; Conservative 31; Mismatches 72; Indels 26; Gaps 8;

QY 25 DWAVTRCFABEAFQAITDNDLPNSLFACNVHOSVFEGEESKEFEG-----LFRTYDDCV 80
 Db 101 NWA-ARKNSEENRDKLT-FEQVFNSTKADNT--SVYVGNISQQTDDADLRLDLSFYGDIA 156
 QY 81 TFOLEKSFRRVRVNFNSPKSAARARIELHETQFRCKKLYFAOVQTPETDGD-----133
 Db 157 EVRIFTQRYAFVRYEKKECATKALMENKEMAGNQCVRGRTQAVVINNQLSNALKI 216
 QY 134 -KLHLP-----PQPAKQFLISPPSPVGVQWOPINDATPVNLNYD---LLYA 175
 Db 217 TVFNLPQNALNPLPIDLSLMMPTMLPTIPLION--PFLNVEPATLLYS 265

RESULT 9
 T42205
 breast cancer susceptibility protein BRCA2 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T42205
 R:McAllister, K.A.; Haugen-Strano, A.; Hagevik, S.; Collins, N.K.; Brownlee, H.; Futr
 submitted to the EMBL Data Library, February 1997
 A:Description: Characterization of the mouse and rat homologs of the BRCA2 breast can
 A:Reference number: Z22073
 A:Accession: T42205
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-3329 <CRA>
A:Cross-references: EMBL:U95652; NID:g2443438; PID:g2443439; PIDN:AAB71377.1
A:Experimental source: strain CD1; 129Sv; ICR Swiss
C:Genetics:
C:Gene: BRCA2
C:Superfamily: breast cancer tumor suppressor BRCA2

Query Watch 7.1%; Score 89; DB 2; Length 3329;
Best Local Similarity 22.0%; Pred. No. 51;
Matches 63; Conservative 30; Mismatches 122; Indels 72; Gaps 12;
QY 6 QQHVPEGGGLFLC---CIDRDWAVTRCFAEAFQATDNDPNSLFACNVHQSVFEG 62
DB 2814 QQVHALQDGAELYAQVYASDPD-HLEACFSEQLRALNNYQMLNDKKQARIQSEFRKA 2872
QY 63 EESKEPFGGLFRDYDCVTFQLEKSFRRVRINFSNPKSAA----RARIETHETQTRGKKL 118
QY 2873 LESAEEKEGLSR--DVTTWKL-----RVTSYKKKEKSALLIWRPSSDLSLLTEGKRY 2925
DB 119 KLYFAQV-----QTPETDGGKLHLAPQAPAKQF--LISPPSS 153
QY 2926 RIYLLAVSKSKSFERPSIQLTAKTRQVQQLPVSSSETLQVYQVPRESLHFSRLSDPAFO 2985
DB 154 PP-----VG-----WOPINDATPVLYND--LLYAVAKLGPGEKYEHLHAGTSTPSVVVH 200
QY 2986 PPCSEVDVGVVSVVRPIGLAPVLYLSDCLNLLVYKFG-----IDLNEDIKPRVLIA 3039
DB 201 VCDSDIEEEDP-----KTSPPKIIQTRRPGLPSPSVSN 234
QY 3040 ASNLQCEPSTSGVPTFLAGHFSIFSAPKEAYFQEKVNNLKHAIEN 3086
DB

RESULT 10
T30904
breast cancer tumor suppressor Brc2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jan-2000
C:Accession: T30904
R:Conor, F.; Smith, A.; Wooster, R.; Stratton, M.; Dixon, A.; Campbell, E.; Tait,
Hum. Mol. Genet. 6, 291-300, 1997
A:Title: Cloning, chromosomal mapping and expression pattern of the mouse Brc2
A:Reference number: Z20931; MUID:97217789; PMID:9063750
A:Accession: T30904
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3329 <CON>
A:Cross-references: EMBL:U82270; NID:g1854950; PID:g1854951; PIDN:AAB48306.1
C:Genetics:
A:Gene: Brc2
A:Map position: 5
C:Superfamily: breast cancer tumor suppressor BRCA2
C:Keywords: tumor suppressor

Query Watch 7.0%; Score 88; DB 2; Length 3329;
Best Local Similarity 22.0%; Pred. No. 63;
Matches 63; Conservative 30; Mismatches 122; Indels 72; Gaps 12;
QY 6 QQHVPEGGGLFLC---CIDRDWAVTRCFAEAFQATDNDPNSLFACNVHQSVFEG 62
DB 2814 QQVHALQDGAELYAQVYASDPD-HLEACFSEQLRALNNYQMLNDKKQARIQSEFRKA 2872
QY 63 EESKEPFGGLFRDYDCVTFQLEKSFRRVRINFSNPKSAA----RARIETHETQTRGKKL 118
QY 2873 LESAEEKEGLSR--DVTTWKL-----RVTSYKKKEKSALLIWRPSSDLSLLTEGKRY 2925
DB 119 KLYFAQV-----QTPETDGGKLHLAPQAPAKQF--LISPPSS 153
QY 2926 RIYLLAVSKSKSFERPSIQLTAKTRQVQQLPVSSSETLQVYQVPRESLHFSRLSDPAFO 2985
DB 154 PP-----VG-----WOPINDATPVLYND--LLYAVAKLGPGEKYEHLHAGTSTPSVVVH 200
QY 2986 PPCSEVDVGVVSVVRPIGLAPVLYLSDCLNLLVYKFG-----IDLNEDIKPRVLIA 3039
DB

1

10

Db 132 ELKAMEELFTSESATFQWLSFRRLRVYDNIATAANARIKLHOYEKNKTVITCVF 191
 QY 123 AQVOTPETDGLHLPQPAKQFLISPPSPVGVQWQINDATPVNLDLLYAVAKLPG 182
 Db 192 AQVTPVSN--KNLOPPAPVKQFLISPPSPAGWEPREGEPLVNDHLLAALASLTPG 248
 QY 183 EYELHAGTSTPSVVHVCDSDIEEDPKTSRPIKIOTRRP 226
 Db 249 ESHELHQSDQPAIIVHTA---MLAETGPGLOVRAPIVOTKCP 289

RESULT 9

CCPL_CAEEL STANDARD; PRT; 207 AA.
 AC P53806: Q9U6V5;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calciopressin-like protein (Down Syndrome candidate region 1-like protein).
 DE RCN-1 OR DSCRIL OR F54E7.7.
 GN Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Petoderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RL Bentley D.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=20221370; PubMed=10756093;
 RA Strippoli P., Lenzi L., Petri M., Carinci P., Zannotti M.;
 ET "A new gene family including DSCR1 (Down syndrome candidate region 1) and Zaki-4: characterization from yeast to human and identification of DSCR1-like 2, a novel human member (DSCRIL2).";
 RT Genomics 64:253-263(2000).
 RL [1]
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses by binding to the catalytic domain of calcineurin A (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U00067; AAK20076.1; --
 DR EMBL; AF176115; AAF01683.1; --
 DR PIR; T34305; T34305.
 DR WormPep; F54E7.7; CE01318
 DR InterPro; IPR006931; Calciopressin.
 DR Pfam; PF04847; Calciopressin; 1.
 FT CONFLICT 205 207
 FT CEQ -> SLQSSSFVLCCTSPSPFFSPFCFNFQSLFCR
 FT G (IN REF. 1)
 SQ SEQUENCE 207 AA; 23030 MW; 0154E308AB05B79 CRC64;

Query Match
 Best Local Similarity 28.0%; Score 351.5; DB 1; Length 207;
 Matches 72; Conservative 38; Mismatches 66; Indels 7; Gaps 4;
 QY 40 ITDFNDLNSLFACNVHVSFEGESKEFGLFTYDCVTFOLFSPFRVRINFSPNK 99
 Db 22 VSSKDDLPNAIVTQVDFVDNFKQDKANFSSLTQTEKDHFDFLSFRVRVIFSSPE 81
 QY 100 SAARIELHETQFGKLLKLYFAQVOTPETDGLHLPQPAKQFLISPPSPVGVQW 159
 Db 82 NATAKLIVQFSFKGHEHKAFFA--ORIYNSANSOMLSPPPLEKQLISPPCSPVGV 139

QY 160 PINDATPVL-NYDLLYAVAKLPGERYELHAGTSTPSVVHVCDSDIEE---EDPKTS 215
 Db 140 QTMDFPVVCFNFDLMARLASFADKEXYEHNGDELTPAIIVHPCETPIDVPSAIEPRT- 198
 QY 216 PKP 218
 Db 199 PRP 201

RESULT 10

YA9F_SCHPO STANDARD; PRT; 163 AA.
 AC Q09791;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C13G6.15c in chromosome I.
 GN SPAC13G6.15c OR SPAC24B11.04C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs K., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leinhardt R., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe".
 CC Nature 415:871-880(2002).
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses by binding to the catalytic domain of calcineurin (By similarity).

CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z54308; CAA91108.1; --
 DR EMBL; Z67757; CAA91769.1; --
 DR PIR; S62444; S62444
 DR GenDB.SPombe; SPAC13G6.15c; --
 DR InterPro; IPR006931; Calciopressin.
 DR Pfam; PF04847; Calciopressin; 1.
 DR Hypothetical protein.

```

171 204      WW.
DOMAIN SEQUENCE 454 AA; 48755 MW; 87CB840D3393EFC0 CRC64;
Query Match      7.8%; Score 98; DB 1; Length 454;
Best Local Similarity 31.1%; Pred. NO. 0.49;
Matches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;
139 PPQAKQFLISPSPSSPPVGQMPQMPINDATPINVDLLYAKLGPGKYELHAGTSTPSV- 197
||||| ||||| ||||| |||||
8 PPQAPQGGQPPSSPPQGGPPS-----GPGQ--PAPATQAQAP 48
198 -----VVHV-CDSDIEEEE-----DPKTSPKPIIOTRRGLPPPS 231
||||| ||||| ||||| |||||
49 PAGHOIVHVRGDSETDLEALFNVMPKTANVQTPMLRLKLPDS 94
db

```

RESULT 12

```

ID      NEOL_RAT      STANDARD      PRPT; 1377 AA.
AC      P97603;
AD      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DT      Neogenin precursor (Fragment).
DE      NEOL OR NGN.
GN      Rattus norvegicus (Rat)..
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC      NCBI_TaxID=10116;
OX      [1]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RC      MEDLINE=97015074; PubMed=8861902;
RX      Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.-Y.,
RA      Culotti J.G., Tessier-Lavigne M.;
RA      Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RT      Cell 87:175-185(1996).
RL      -1- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC      TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC      DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC      MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC      SUBFAMILY.
CC      -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 6 fibronectin type III domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement. See http://www.isb-slb.ch/announce/
CC      -----

```

```

CC      or send an email to license@ncbi.nlm.nih.gov
CC      -----
DR      EMBL: U68726; AAB41100.1; -.
DR      HSSP: P56276; 1TLK.
DR      InterPro: IPR003961; FN.III.
DR      InterPro: IPR003962; FN.II_subod.
DR      InterPro: IPR007110; Ig-like.
DR      InterPro: IPR003598; Ig_C2.
DR      InterPro: IPR003006; Ig_MHC.
DR      Pfam: PF00041; fn3; 6.
DR      Pfam: PF00047; Ig; 4.
DR      PRINTS: PR00014; FNTYPEIII.
DR      SMART: SM00060; FN3; 6.
DR      SMART: SM00408; IGC2; 3.
DR      PROSITE: PS50833; IG_LIKE; 4.
DR      Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;
KW      Glycoprotein.
KW      NON_TER
KW      POTENTIAL.
FT      SIGNAL
FT      CHAIN
FT      1
FT      2
FT      3 1377
FT      CHAIN

```

[illegible]

RESULT 11	YAP1_HUMAN	STANDARD;	PRT;	454 AA.
ID	YAP1_HUMAN			
AC	P46937;			
DC	01-NOV-1995 (Rel. 32, Created)			
DD	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-SEP-2003 (Rel. 42, Last annotation update)			
DR	65 kDa Yes-associated protein (YAP65).			
DE	YAP1 OR YAP65.			
GN	Homo sapiens (Human).			
OS	Chordata; Vertebrata; Euteleostomi;			
OC	Eumariota; Metazoa;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;			
PC	TMEDLINE=95301570; PubMed=7782339;			
RX	Sudol M., Bork P., Einbond A., Kastury K., Druck T., Negrini M.,			
RA	Huebner K., Lehman D.;			
RT	"Characterization of the mammalian YAP (Yes-associated protein) gene			
RT	and its role in defining a novel protein module, the WW domain.";			
FT	270-14733-14741(1995).			

J. Biol. Chem. 270:1493-1498, 1995.
[2] INTERACTION WITH WBP1 and WBP2.
MEDLINE=97347517; PubMed=9202023;
Chen H.-I., Einbond A., Kwak S.-J., Linn H., Koepf E., Peterson S.,
RA Kelly J.W., Sudol M.;
ET Characterization of the WW domain of human Yes-associated protein and
its polyproline containing ligands.;
PT J. Biol. Chem. 272:17070-17077(1997).
RL J. Biol. Chem. 272:17070-17077(1997).
CC -1- SUBUNIT: Binds to the SH3 domain of the YES kinase. Binds to WBP1
and WBP2.
CC -2- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
CC -3- SIMILARITY: Contains 1 WW domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X80507; CNA56672.1; -.
CC PIR; A56954; A56954.
CC PDB; 1JMO; 21-DEC-01.
CC PDB; 1K9Q; 28-DEC-01.
CC PDB; 1K9R; 28-DEC-01.
CC Genew; HGNC:16262; YAP1.
CC MIM; 606608; -.
CC GO; GO:0005515; F:protein binding activity; TAS.
CC InterPro; IPR005153; Mbth.
CC InterPro; IPR001202; WW_Rsp5_WWP.
CC Pfam; PF00397; WW; 1.
CC SMART; SM00456; WW; 1.
CC PROSITE; PS01159; WW_DOMAIN_1; 1.
CC PROSITE; PS00020; WW_DOMAIN_2; 1.
CC Phosphorylation; 3D-structure.
KW

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 18:58:43 ; Search time 60 Seconds
(without alignments)
514.651 Million cell updates/sec

Title: US-10-030-613-1

Perfect score: 1254

Sequence: 1 MRSPPQOQHVPEDGGLFLLC.....SPKPKIIOTRRPGLPPSVSN 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	913	72.8	192	10	US-09-782-953-18
2	913	72.8	192	10	US-09-782-953-19
3	905	72.2	197	10	US-09-782-953-9
4	905	72.2	197	10	US-09-782-953-10
5	729.5	58.2	241	10	US-09-782-953-21
6	729.5	58.2	241	10	US-09-782-953-22
7	698.5	55.7	212	10	US-09-782-953-24
8	698.5	55.7	212	10	US-09-782-953-25
9	629	50.2	197	10	US-09-782-953-12
10	629	50.2	197	10	US-09-782-953-13
11	629	50.2	197	10	US-09-782-953-15
12	629	50.2	197	10	US-09-782-953-16
13	623	49.7	198	10	US-09-782-953-6
14	623	49.7	198	10	US-09-782-953-7
15	620.5	49.5	198	10	US-09-782-953-3

SUMMARIES

16	630.5	49.5	198	10	US-09-782-953-4	Sequence 4, Appli
17	282	22.5	142	9	US-09-925-302-790	Sequence 790, App
18	222	17.7	58	9	US-09-864-761-35379	Sequence 35379, A
19	222	17.7	58	9	US-09-864-761-43076	Sequence 43076, A
20	171	13.6	56	9	US-09-864-761-34111	Sequence 34111, A
21	112	8.9	32	15	US-10-194-155-8	Sequence 8, Appli
22	90	7.2	4019	10	US-09-738-973-425	Sequence 425, App
23	90	7.2	4019	10	US-09-854-133-425	Sequence 425, App
24	90	7.2	4019	10	US-10-144-649A-425	Sequence 425, App
25	84.5	6.7	212	15	US-10-278-173-122	Sequence 122, App
26	84.5	6.7	212	15	US-10-295-403-16	Sequence 16, Appl
27	84	6.7	3063	15	US-10-177-293-61	Sequence 63, Appl
28	83.5	6.7	229	11	US-09-934-455-508	Sequence 508, App
29	83.5	6.7	434	15	US-10-102-806-506	Sequence 506, App
30	83.5	6.7	659	15	US-10-211-962-12	Sequence 12, Appl
31	83.5	6.7	3063	12	US-10-301-822-26	Sequence 26, Appl
32	83.5	6.7	3063	15	US-10-177-293-61	Sequence 61, Appl
33	82	6.5	526	10	US-09-801-368-362	Sequence 362, App
34	81.5	6.5	916	10	US-09-919-497-64	Sequence 1610, Ap
35	81.5	6.5	916	10	US-09-925-300-1610	Sequence 10, Appl
36	80.5	6.4	296	10	US-09-789-054A-10	Sequence 4, Appli
37	80.5	6.4	932	10	US-09-754-997A-2	Sequence 21, Appl
38	80.5	6.4	1252	10	US-09-908-193-21	Sequence 83, Appl
39	80.5	6.4	2701	15	US-10-171-311-83	Sequence 2, Appli
40	80.5	6.3	1706	15	US-10-024-450-2	Sequence 3, Appli
41	79.5	6.3	1706	15	US-10-142-650-3	Sequence 7, Appli
42	79.5	6.3	434	15	US-10-192-634-7	Sequence 7, Appli
43	79	6.3	434	15	US-09-854-549-7	Sequence 6, Appli
44	79	6.3	559	9	US-09-836-561-6	
45	79	6.3	559	9		

ALIGNMENTS

RESULT 1

US-09-782-953-18
; Sequence 18, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: CALCIUM-INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSID:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-18

Query Match	72.8%	Score	913	DB	10	Length	192
Best Local Similarity	93.0%	Pred. No.	4.8e-84				
Matches	174	Conservative	3	Mismatches	10	Indels	0
Gaps	0						
QY	48	NSLFACNVHQSVEGEEKPEGLFRYDDCVTFQLEKSFRRVIRNFSNPKSAARIE	107				
Db	6	STLVACVVDVEFTTQEVKEKFGGLFRYDDCVTFQLEKSFRRVIRNFSNPKSAARIE	65				
QY	108	LHETFRGKLLKLYFAQVOTPETDGDKLHLAPPAPAKQFLISPPSPVGVQNPNDATPV	167				
Db	66	LHETFRGKLLKLYFAQVOTPETDGDKLHLAPPAPAKQFLISPPSPVGVQNPNDATPV	125				
QY	168	LNVDLLYAVAKLGKGYELHAGTSTPVSVVHVHCDSDIEEDPKTSPKPKIIQTRPG	227				
Db	126	LNVDLLYAVAKLGKGYELHAGTSTPVSVVHVHCDSDIEEDPKTSPKPKIIQTRPG	185				

```

48  NSLFACNVHQSVEGBESKEFEGFRTYDDCVTFQLKSFRRVRINFSPKSAARIE 107
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
11  STLVCVVDVEVTNOEVKEKPEGLFRTYDECVTFQLKSFRRVRINFSPKSAARIE 70
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
108 LHETQFRGKKLKLYFAQVOTPTDGDKHLAPQAKQFLISPPSPVGVQPIINDATPV 167
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
71  LHETQFRGKKLKLYFAQVOTPTDGDKHLAPQAKQFLISPPSPVGVQNPISDATPV 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
168 LNYDLLYAVAKLGEGKEYELHAGTESTPSVVHVCDSDIEEEDPKTSPKPIIOTRRPG 227
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
131 LNYDLLYAVAKLGEGKEYELHAGTESTPSVVHVCDSDMEEEDPKTSPKPIIOTRRPG 190
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
228 LPSPVSN 234
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
191 LPSPVSN 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 4

```

RESULT 4
US-09-782-953-10
; Sequence 10, Application US/09782953
; Patent No. US2002015093A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent.in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-10

```

[illegible]

RESULT 5

```

RESULT 5
US-09-782-953-21
; Sequence 21 Application US/09782953
; Patent NO. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCIINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE NUMBER: US/09/782.953
; CURRENT APPLICATION NUMBER: 2001-02-13

```

```

228 LPPSVN 234
111111
186 LPPSVN 192

Db

RESULT 2
US-09-782-953-19
; Sequence 19, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

```

[illegible]

RESULT 3

```

US-09-782-953-9
; Sequence 9, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-782-953-9

```

```

Query Match          72.2%;   Score 905;   DB 10;   Length 197;
Best Local Similarity 91.4%;   Pred. No. 3.2e-83;
Matches 171; Conservative      8;   Mismatches 8;   Indels 0;   Gaps 0;

```


Result No.	Query Match	Score	Query			Description	
			Match	Length	DB		
1	1	913	72.8	197	1	CCP2_HUMAN	Q14206 homo sapien
2	2	905	72.2	197	1	CCP2_MOUSE	Q9jhg2 mus musculus
3	4	733.5	58.5	239	1	CCP3_MOUSE	Q9jkk0 mus musculus
4	3	729.5	58.2	241	1	CCP3_HUMAN	Q9uka8 homo sapien
5	5	632	50.4	197	1	CCP1_CRTR	Q95847 cricetus
6	6	629	50.2	197	1	CCP1_HUMAN	P53805 homo sapien
7	7	622.5	49.6	198	1	CCP1_MOUSE	Q9jhg6 mus musculus
8	9	432.5	34.5	292	1	SRA_DROME	Q9xz18 drosophila
9	9	351.5	28.0	207	1	CCPL_CAMEL	P53806 caenorhabdi
10	10	107.5	8.6	163	1	YAP1_SCHPO	Q09791 schizosacch
11	11	98	7.8	454	1	YAP1_HUMAN	P46937 homo sapien
12	12	97.5	7.8	1377	1	NEOL_RAT	P97603 rattus norv
13	13	89	7.1	3726	1	ABF1_MOUSE	Q61329 mus musculus
14	14	88	7.0	3329	1	BRC1_MOUSE	P97929 mus musculus
15	15	86.5	6.9	1331	1	MANB_CALSA	P22533 caldocellum
16	16	86	6.9	668	1	PAB5_ARATH	Q03196 arabidopsis
17	17	85	6.8	1232	1	YOO5_CAMEL	P34643 caenorhabdi
18	18	84	6.7	1206	1	PM14_MOUSE	Q05859 mus musculus
19	19	84	6.7	1468	1	FMN1_MOUSE	Q05860 mus musculus
20	20	83.5	6.7	3063	1	CA1C_HUMAN	Q93715 homo sapien
21	21	83	6.6	408	1	IF13_MOUSE	Q35168 mus musculus
22	22	83	6.6	853	1	OTU_DROME	P10383 drosophila
23	23	83	6.6	5120	1	PCLO_CHICK	Q9pu36 gallus gall
24	24	82.5	6.6	457	1	MUC_SUNMU	P20768 suncus muri
25	25	82.5	6.6	663	1	MX1_PIG	P27594 sus scrofa
26	26	82	6.5	425	1	YTR1_EBV	P30119 epstein-bar
27	27	82	6.5	434	1	BC13_MOUSE	P59017 mus musculus
28	28	82	6.5	526	1	CAP_YEAST	P17555 saccharomyc
29	29	81.5	6.5	905	1	XPF_HUMAN	Q92889 homo sapien
30	30	81.5	6.5	2319	1	NTG3_RAT	Q9rl72 rattus norv
31	31	81	6.5	639	1	CA1C_RABIT	Q28902 orthotolagus
32	32	80.5	6.4	587	1	SYD_SFRCO	Q9f323 streptomyces
33	33	80.5	6.4	611	1	PE54_YEAST	P39684 saccharomyce

DR EMBL; AK019377; BAB31587.1; -
 DR MGD; MG1:1858220; DSCR1L2.
 DR InterPro: IPR006931; Calcipressin.
 DR Pfam: PF04847; Calcipressin; 1.
 SQ SEQUENCE 239 AA; 27153 MW; 1B2687B47BB4D272 CRC64;
 Query Match 58.5%; Score 733.5; DB 1; Length 239;
 Best Local Similarity 67.1%; Pred. No. 6.6e-56;
 Matches 143; Conservative 24; Mismatches 43; Indels 3; Gaps 2;
 QY 19 LCICDRDWAIVTRCFA--EEAFQAITDFNDLNSFACNVHQSVFEGESKEKEGLFRTY 76
 DB 16 LCSSDQEEEMVFNEDGEGLEMDLSDPTSLFACSVHEAFVQEQERFEALFTLY 75
 QY 77 DCVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136
 DB 76 DDQVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136
 QY 137 LAPPOPAKQFLISPPSPVGVQPIINDATPVNLDLYAVAKLGPGKGYELHAGTESTPS 196
 DB 136 LLPQPTKQFLISPPSPVGVQKQSEDAMPVINDLLCAVSKLGPGKGYELHAGTESTPS 195
 QY 197 VVHVCDSDIEEDPKTSPKPIIOTRRPGLP 229
 DB 196 VVHVCDSETEEDTK-NPKQKITOTRRPEAP 227

RESULT 4

CCP3_HUMAN
 ID CCP3_HUMAN STANDARD; PRT; 241 AA.
 AC Q9UKA8; Q9UKA7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
 DE {Myocyte-enriched calcineurin interacting protein 3} (MCIP3).
 GN DSCR1L2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Peripheral blood, and placenta;
 RX MEDLINE=2021370; PubMed=10756093;
 RA Strippoli P., Lenzi L., Petrini M., Carinci P., Zannotti M.;
 RT "A new gene family including DSCR1 (Down syndrome candidate region 1)
 and ZAKI-4: characterization from yeast to human and identification of
 DSCR1-like 2, a novel human member (DSCR1L2).";
 RL Genomics 64:252-263(2000).
 RN [2]
 RP SEQUENCE OF 66-241 FROM N.A.
 RA Bagguley C.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
 by binding to the catalytic domain of calcineurin A. Could play a
 role during central nervous system development (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9UKA8-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9UKA8-2; Sequence=VSP_001319;
 CC -1- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
 kidney, liver and peripheral blood leukocytes. Lower expression in
 all other tissues.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announcement/
 or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF176116; AAF01684.1; -
 DR EMBL; AF176117; AAF01685.1; -
 DR EMBL; AL034582; CAB72298.1; -
 DR Genew; HGNC:3042; DSCR1L2.
 DR MIM: 605860; -
 DR GO: GO:0003723; F:RNA binding activity; TAS.
 DR GO: GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro: IPR006931; Calcipressin.
 DR Pfam: PF04847; Calcipressin; 1.
 DR Alternative splicing.
 KW VARSPLIC 124 133 Missing (in isoform 2).
 FT VARSPLIC 124 133 /FTID=VSP_001319.
 SQ SEQUENCE 241 AA; 27492 MW; 91AB619F5E443FBD CRC64;

Query Match 58.2%; Score 729.5; DB 1; Length 241;
 Best Local Similarity 66.1%; Pred. No. 1.5e-55;
 Matches 144; Conservative 22; Mismatches 49; Indels 3; Gaps 2;
 QY 19 LCICDRDWAIVTRCFA--EEAFQAITDFNDLNSFACNVHQSVFEGESKEKEGLFRTY 76
 DB 16 LCSTDQEEEMIFGENEDDLDMDLSDPTSLFACSVHEAFVQEQERFEALFTY 75
 QY 77 DCVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136
 DB 76 DDQVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136
 QY 137 LAPPOPAKQFLISPPSPVGVQPIINDATPVNLDLYAVAKLGPGKGYELHAGTESTPS 196
 DB 136 LLPQPTKQFLISPPSPVGVQKQSEDAMPVINDLLCAVSKLGPGKGYELHAGTESTPS 195
 QY 197 VVHVCDSDIEEDPKTSPKPIIOTRRPGLP 234
 DB 196 VVHVCDSETEEDTK-NPKQKITOTRRPGLP 232

RESULT 5

CCP1_CRIGR
 ID CCP1_CRIGR STANDARD; PRT; 197 AA.
 AC Q35847;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
 DE (Oxidative-induced protein Adapt78).
 GN DSCR1 OR ADAPT78.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97329095; PubMed=9185608;
 RT "Hamster adapt78 mRNA is a Down syndrome critical region homologue
 that is inducible by oxidative stress.";
 RL Arch. Biochem. Biophys. 342:6-12(1997).
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
 by binding to the catalytic domain of calcineurin A. Could play a
 role during central nervous system development (By similarity).
 CC -1- INDUCTION: OXIDANT-INDUCIBLE.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announcement/
 or send an email to license@isb-sib.ch).

Tue Aug 26 09:11:04 2003

```

CC -----
DR EMBL: U60263; AAB68517.1; -
DR InterPro: IPR006931; Calcipressin.
DR Pfam: PF04847; Calcipressin; 1.
SQ SEQUENCE 197 AA; 22674 MW; 0EFC68FA55826439 CRC64;

Query Match          50.4%; Score 632; DB 1; Length 197;
Best Local Similarity 63.7%; Pred. NO. 2.6e-47;
Matches 123; Conservative 24; Mismatches 38; Indels 8; Gaps 2;

QY 42 DENDLPSNLFACNVHVSFEGEESKEFGLFRYDCTVTFQLFKFRVRVNFSPKSA 101
DB 5 DFNVSFLLACVANGDVFSESETRAKFSLFRYDKDITQYFKFRVRVNFSPKSA 64
QY 102 ARAIEHEHETFRGKKLKLKLYFAQVQTPETDGLHLAPPQAPAKOFLISPPSPVGHQPI 161
DB 102 ARAIEHEHETFRGKKLKLKLYFAQVQTPETDGLHLAPPQAPAKOFLISPPSPVGHQPI 161
QY 65 ADARLQHLKTEFLGKMKLYFAQTLHGSS-----HLAPPNPDKOFLISPPSPVGHQV 120
DB 65 ADARLQHLKTEFLGKMKLYFAQTLHGSS-----HLAPPNPDKOFLISPPSPVGHQV 120
QY 162 NDATPVLNVDLLYAVAKLGPGEKYLHAGTSTSVVHVCDSDIEEDPKTS-----PK 217
DB 162 NDATPVLNVDLLYAVAKLGPGEKYLHAGTSTSVVHVCDSDIEEDPKTS-----PK 217
QY 121 EDATPVINYDLYAISKLGPKGYELHAATDTTSPVVHVCDSDIEEDPKTS-----PK 180
DB 121 EDATPVINYDLYAISKLGPKGYELHAATDTTSPVVHVCDSDIEEDPKTS-----PK 180
QY 218 PKIOTRRRGLPP 230
DB 181 PKIOTRRRPEYTP 193

RESULT 6
CCPI_HUMAN STANDARD: PRT; 197 AA.
ID AC P53805; O00582; O00583; Q96R03; Q9B069; Q9UF15; Q9UME4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1) (Myocyte-
DE enriched calcineurin interacting protein 1) (MCIP1) (Adapt78).
DE DSCR1 OR DSC1 OR ADAPT78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
[1]
RP MEDLINE=96121593; PubMed=8955418;
RX Fuentes J.-J., Pritchard M.A., Planas A.M., Bosch A., Ferrer I.,
RA Estivill X.;
RT "A new human gene from the Down syndrome critical region encodes a
RT proline-rich protein highly expressed in fetal brain and heart.";
RL Hum. Mol. Genet. 4:1935-1944(1995).
[2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP MEDLINE=97468152; PubMed=9325050;
RX Fuentes J.-J., Pritchard M.A., Estivill X.;
RA "Genomic organization, alternative splicing, and expression patterns
of the DSCR1 (Down syndrome candidate region 1) gene.";
RL Genomics 44:358-361(1997).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 3).
RP Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,
RA Qiang B.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP TISSUE=Lung;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Donaldson M.F., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

```

```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Small D.E.,
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Marra M.A.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RN SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).
RP TISSUE=Breast; Leahy K.P., Davies K.J.A.;
RC Crawford D.R., a calcium and oxidant-inducible RNA.";
RA "Adapt78, a calcium and oxidant-inducible RNA.";
RT Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[6]
RN CHARACTERIZATION.
RP MEDLINE=20320698; PubMed=10861295;
RX Fuentes J.J., Genesca L., Kingsbury T.J., Cunningham K.W.,
RA Perez-Riba M., Estivill X., de la Luna S.;
RA "DSCR1, overexpressed in Down syndrome, is an inhibitor of
RA calcineurin-mediated signaling pathways.";
RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of
RT calcineurin-mediated signaling pathways.";
RL Hum. Mol. Genet. 9:1681-1690(2000).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=P53805-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P53805-2; Sequence=VSP_001314;
CC Name=3;
CC IsoId=P53805-3; Sequence=VSP_001315;
CC Name=4;
CC IsoId=P53805-4; Sequence=VSP_001316;
CC -!- TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal
CC muscle. Also expressed in all other tissues.
CC -!- INDUCTION: By calcium.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U28833; AAB81557.1; -
CC EMBL: U85265; AAB84370.1; -
CC EMBL: U85266; AAB84371.2; -
CC EMBL: U85267; AAB84372.1; -
CC EMBL: AF400429; AAK92478.1; -
CC EMBL: BC002864; AAH02864.1; -
CC EMBL: U53821; AAF21218.1; -
CC Genew; HGNC:3040; DSCR1.
CC MIM; 602917; -
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0003700; F:transcription factor activity; TAS.
CC GO; GO:0007417; P:central nervous system development; TAS.
CC GO; GO:0008015; P:circulation; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro: IPR006931; Calcipressin.
CC Pfam: PF04847; Calcipressin; 1.
CC Alternative splicing.
CC DOMAIN 166 174
CC VARSPLIC 1 28
CC POLY-GLU.
CC MEEVDQLQPSATIAChLPRVVDGLC -> MHERNENYS
CC FSSLIACVANSIDFSESET (in isoform 2).
CC /FTId=VSP_001314.

```

C:Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match
Best Local Similarity 6.9%; Score 87; DB 2; Length 671;
Matches 50; Conservative 37; Mismatches 94; Indels 56; Gaps 12;
QY 23 DRWAVTRC-----FAEEAFOAITDFNDLP--NSLFACNVHQSVEGEEKFKFG 71
DB 291 DKEWFGKAOKSRETELKQFEQSLKEAADSQGSNLYVKNLDSVTD-DKLRHF-A 348
QY 72 LFRIYDVCVTFQLEKSPRRVR--INFSPKSAARARIELHETQFGKLL---KLFAQVQ 126
DB 349 PFGTITCKVMRDPGSGVSGFVAFSTPEATRA-----ITEMGKMIVTKPLVALAQ 403
QY 127 TPETDGDKLHL-----APPQAKQFLISPPSPVVGW-----PINDAPPVLNY 170
DB 404 RKEDRKARLQAQFSQMRPNMPPAVGPRMOMYPPGPGMGQQLFYGGPPAMIPQPGFY 463
QY 171 DLLYAVAKLPGKEKYLHAGTSTPVSVVHVCDSDIEEDPKTSPKIIOTRRPG 227
DB 464 Q-----QQLVPG-----MRPGGSPMPNFFPMNQOQOQOQOQO-----QOORPG 503

RESULT 13

A48954
mannan endo-1,4-beta-mannosidase (EC 3.2.1.78) - Caldocellum saccharolyticum
N:Alternate names: beta-mannanase
C:Species: Caldocellum saccharolyticum
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: A48954; B43745
R:Gibbs, M.D.; Saul, D.J.; Luthi, E.J.; Bergquist, P.L.
Appl. Environ. Microbiol. 58, 3864-3867, 1992
A:Title: The beta-mannanase from "Caldocellum saccharolyticum" is part of a multidomain
A:Reference number: A48954; MUID:93119139; PMID:1476429
A:Accession: A48954
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-1331 <GIB>

A:CROSS-references: GB:L01257; NID:g144290; PIDN:AAAY1887.1; PID:g144291
R:Luthi, E.; Bhana Jasmal, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
Appl. Environ. Microbiol. 57, 694-700, 1991
A:Title: Cloning, sequence analysis, and expression in *Escherichia coli* of a gene coding
A:Reference number: A43745; MUID:91247819; PMID:2039230
A:Accession: B43745
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337, 'PPRQHQHRO' <LUE>

A:CROSS-references: EMBL:M36063; NID:g144292; PIDN:AAAY2861.1; PID:g144294
A:Note: the authors translated the codon CAC for residue 262 as Glu
A:Note: this sequence has been revised in reference A48954
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match
Best Local Similarity 6.9%; Score 86.5; DB 2; Length 1331;
Matches 50; Conservative 28; Mismatches 67; Indels 79; Gaps 10;
QY 24 RDWAVTRCFAEFAFOAITDFNDLPNSLFACNVHQSVEGEEKFKFGFLRTYDDCVTFQ 83
DB 201 QDMSNT---MRDQAQSIEMADPLRLNVFSIHMY-----GVYNT-----235
QY 84 LFKSPRRVRINFSPKSAARARIELHETQFGKLLKLYFAOVQTPETDGDKLHLAPPOA 143
DB 236 -----ASKVEYIKSFVDKGLPLVIGFEGHQHTDGDPOEATVRYA 276
QY 144 KQFLISPPS-----SPVVG-----WQPINATPVLYNLDLYAVAKLPGKEKYLHA- 189
DB 277 KQYKLGFSWCSWCSNYSYGLDMVNNWDP--NNPTW-----GWYKTNAI 321
QY 190 GTESTPVSVVHVCDSDIEEDPKTSPKIIOTRRPGLPPSVS 233
DB 322 GTSSTPTPT-----SVVTPPTPTPTPTPTVATPTT--TPTPTVS 359

RESULT 14

B96740
hypothetical protein F14023.15 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: B96740
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Altschuld, R.; Holt, D.; Lander, E.S.; Zeng, K.; Shinn, P.; Shinn, P.; Southwick, A.M.; Sun, H.; Talbot, W.H.; Wehner, J.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
Nature 408, 815-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, R.; Rizzio, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talbot, W.H.; Wehner, J.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B96740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-668 <STO>

A:CROSS-references: GB:AE005173; NID:g7239504; PIDN:AAF43230.1; GSPDB:GN00141
C:Genetics:
A:Gene: F14023.15
A:Map position: 1
C:Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match
Best Local Similarity 6.9%; Score 86; DB 2; Length 668;
Matches 37; Conservative 28; Mismatches 56; Indels 36; Gaps 7;
QY 23 DRWAVTRCFAEFAFOAITDFNDLPNS-LFACNVHQSVEGEEKFKFGFLRTYDDCVT 81
DB 305 DREELARKFEQ---RISRFKQLQGSNLYKLNLDSDV-----NDEKLEMESEYGNVTS 356
QY 82 FOLPKSPRRVR-----INFSPKSAARARIELH-----ETQFGKLLKLY 121
DB 357 CKVMNSQGLSRGFGFVAYSNPSEALLANKMGKMGKPLLYVALAQRKEERQALQSL 416
QY 122 FAOVQTPETDGDKLHLAP-POPQKQLISPPSPVVG 157
DB 417 FTQIRSPGT-----MSPVPSMSPGHHHPGPGMSG 447

RESULT 15

S40766
hypothetical protein ZK512.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Sep-1997
C:Accession: S40766
R:Hawkins, T.; Ainscough, R.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40759
A:Accession: S40766
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1232 <HAW>
A:CROSS-references: ENBL:222177; NID:g297989; PID:g297994
A:Introns: 5/3; 19/1; 52/2; 108/2; 306/3; 344/2; 490/3; 542/3; 634/3; 703/2; 1060/3;

Query Match
Best Local Similarity 6.8%; Score 85; DB 2; Length 1232;
Matches 50; Conservative 32; Mismatches 78; Indels 72; Gaps 9;

QY 27 AVTRCFEAEAF-----QAITD--FNDLPNS-----LFACNV 55
DB 212 SVTNTLASEVNSSESYGVVVKLGSALVDHVTMPPEKYYKMPIDRVVYFYCAV 271
QY 56 HQSVFEGEEKFKFGL-----RTYDDCVTFQFKSFRFRVINFSPKSA-----102
DB 272 NNKPY---NNIDFHVIFNREFSYRGYSGDKDLALFKVKRMQEFSLKQLEADRLAYE 328

us-10-030-613-1-1.rpr

Tue Aug 26 09:11:03 2003

Qy	103	RARIELHETQFRGKKLYFAQVOTPDGDKLHLPQPAKQFLISPPSPVGMOPIN	162
Db	329	KARQEAASE-----KLDNFQHKIEEREPEKLNISQPE---EVLNNGP-----	368
Qy	163	DATPVLNYDLLYAVAKLGPEKEYELHAGTESTPSVVVHVCDSDIEEEDPKT	214
Db	369	-----LHYHSLQFATIGVGGKLVIIKPAGTIDTSITGHVLTSSVHVDDLKT	415

Search completed: August 22, 2003, 18:59:20
Job time : 44 secs

[illegible]

Tue Aug 26 09:11:02 2003

US-09-782-953-3

Query Match	49.5%	Score	620.5	DB	10	Length	198
Best Local Similarity	62.4%	Pred. No.	1.5e-54				
Matches	121	Conservative	22	Mismatches	42	Indels	9
						Gaps	2

QY	42	DENDLPNSLFACNVHQSVFEGESKEKEFGLFRITYDDCVTFOLPKSFRVRINFSNPKSA	101
Db	5	DIQDLPSATIACHLDPRVFDGCLCRAKTESLFRITYDKDTTFQYFKSKRRVRINFSNPLSA	64
QY	102	ARARIELHETQPRGKKLKLYFAOVQTPETDGDKHLAPPQPAKQFLISPPSSPPVGVQPI	161
Db	65	ADARLRLHKTFLGKEMKLYFAOTLHIGSS----HLAPPNDKQLISPPASPPVGVKQV	120
QY	162	NDATPVNLVLLYAVAKLGPGEKYLHAGTESTPSVVHVHVCDSIEEEDPK-----TSP	216
Db	121	EDATPVINYDLLYAIKSLGPEKYLHAATDPTPSVVHVHVCESDQENEEEMERMRKP	180
QY	217	KPKIOTRRRGLPP	230
Db	181	KPKIOTRRPEYTP	194

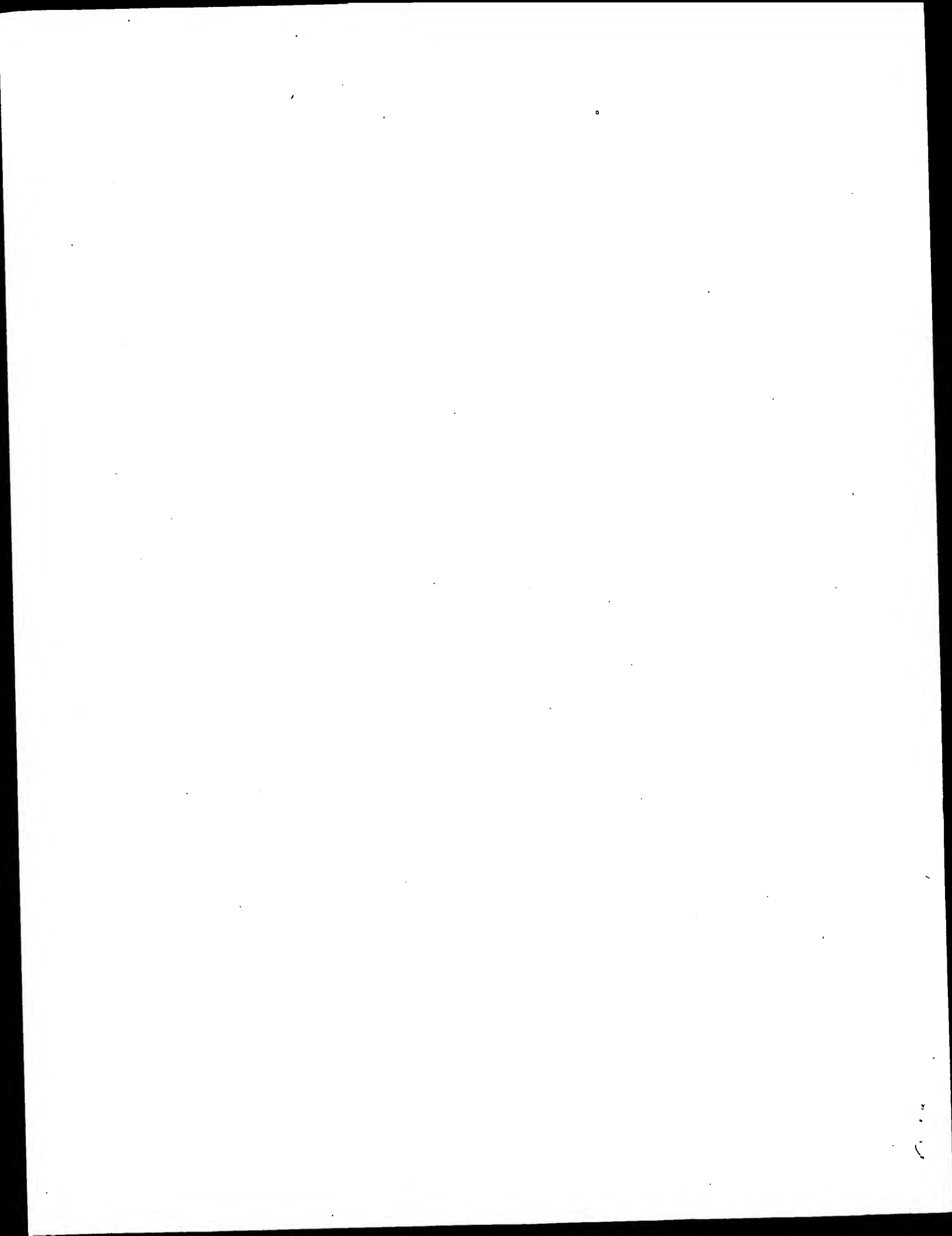
Search completed: August 22, 2003, 19:08:07
Job time : 61 secs

Query Match	7.1%; Score 89; DB 1; Length 3726;
Best Local Similarity	28.5%; Pred. No. 41;
Matches	43; Conservative 17; Mismatches 65; Indels 26; Gaps 7;
QY	29 TRCFABE-AFOAITDFNDLPNSLFACNVHQSVEGESEKKEFGLFRYDDCVTFQFLPKS 87
DB	1951 TRALLEFNGFELVIOINE--NKQAKKNGKAEQGESLEKLE-----CDSC-----CKL 1998
QY	88 FRVRVNFNFKSAAR-----ARIELHETQFRGKKLYFAQVQPP-----TDGDKL 135
DB	1999-FSNILLKSHQEHVQNYFPFQLEKRFKQYREHYDKLYPLRQNPPEPPPPPPPL 2058
QY	136 HIAPPQAKOFLISPPSPVQWQINDATP 166
DB	2059 PTAPPQAPAPAI-PASAPPITSPITAPAP 2088
RESULT 14	
BRC2_MOUSE	
ID	BRC2_MOUSE STANDARD; PRT; 3329 AA.
AC	P97929; O35922; P97383;
DT	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DE	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Breast cancer type 2 susceptibility protein.
GN	BRC2.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC	NCBI_TaxID=10090;
EN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=129; SV;
RC	MEDLINE=97317789; PubMed=9063750;
RA	Connor F., Smith A., Wooster R., Stratton M., Dixon A., Campbell E.,
RA	Tait T.M., Freeman T., Ashworth A.;
RA	"Cloning, chromosomal mapping and expression pattern of the mouse
RA	Brc2 gene.";
RL	Hum. Mol. Genet. 6:291-300(1997).
EN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6;
RC	MEDLINE=9737041; PubMed=9119389;
RA	Sharan S.K., Bradley A.;
RA	"Murine Brc2: sequence, map position, and expression pattern.";
RL	Genomics 40:234-241(1997).
EN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=129/SV;
RC	MEDLINE=97384941; PubMed=9242436;
RA	McAllister K.A., Haugen-Strano A., Hagevik S., Brownlee H.A.,
RA	Collins N.K., Futreal P.A., Bennett L.M., Wiseman R.W.;
RA	"Characterization of the rat and mouse homologues of the BRC2 breast
RA	cancer susceptibility gene.";
RL	Cancer Res. 57:3121-3125(1997).
EN	[4]
RP	SEQUENCE OF 18-200 FROM N.A.
RC	MEDLINE=97075121; PubMed=9917547;
RA	Rajan J.V., Wang M., Marquis S.T., Chodosh L.A.;
RA	"Brc2 is coordinately regulated with Brcal during proliferation and
RA	differentiation in mammary epithelial cells.";
RL	Proc. Natl. Acad. Sci. U.S.A. 93:13078-13083(1996).
EN	[5]
RP	SEQUENCE OF 569-625 FROM N.A.
RC	MEDLINE=97341126; PubMed=9196008;
RA	McAllister K.A., Ramachandran S., Haugen-Strano A., Fiedorek F.T. Jr.,
RA	Wiseman R.W.;
RA	"Genetic mapping of the Brc2 breast cancer susceptibility gene on
RA	mouse chromosome 5.";
RL	Mamm. Genome 8:540-541(1997).
CC	-1- FUNCTION: MAY PARTICIPATE IN A PATHWAY ASSOCIATED WITH THE
CC	ACTIVATION OF DOUBLE-STRAND BREAK REPAIR AND/OR HOMOLOGOUS
CC	RECOMBINATION (BY SIMILARITY).

-1- SUBUNIT: INTERACTS WITH RAD51 (BY SIMILARITY).
-1- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGHEST EXPRESSION IN
CEREBELLUM, TESTIS, ILEUM, APPENDIX, EPIDYMI, OVARY AND MAMMARY
GLAND. NO EXPRESSION IN LONG.
-1- DEVELOPMENTAL STAGE: IN THE MAMMARY GLAND, EXPRESSION INCREASES
DRAMATICALLY DURING PREGNANCY.
-1- SIMILARITY: Contains 7 BRC2 repeats.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; U82270; AAB48306.1; -
EMBL; U72947; AAB40720.1; -
EMBL; U65594; AAC23702.1; -
EMBL; U89652; AAB71377.1; -
EMBL; U89503; AAC53276.1; -
PIR; T30904; T30904.
PDB; 1MIU; 25-SEP-02.
MGD; MGI:109337; Brc2.
GO; GO:0005737; C:cytoplasm; IDA.
GO; GO:0005634; C:nucleus; ISS.
GO; GO:0005716; C:synaptonemal complex; ISS.
GO; GO:0005515; F:protein binding activity; ISS.
GO; GO:0003697; F:single-stranded DNA binding activity; ISS.
GO; GO:0003653; F:transcriptional activator activity; ISS.
GO; GO:0006915; P:apoptosis; ISS.
GO; GO:0006338; P:chromatin modeling; ISS.
GO; GO:0006260; P:DNA replication; ISS.
GO; GO:0000724; P:double-strand break repair via homologous r. .; ISS.
GO; GO:0006325; P:establishment and/or maintenance of chromatin; ISS.
GO; GO:0007093; P:mitotic checkpoint; ISS.
GO; GO:0007090; P:regulation of S phase of mitotic cell cycle; ISS.
GO; GO:0045449; P:regulation of transcription; ISS.
InterPro; IPR002093; BRC2_repeat.
Pfam; PF00634; BRC2; 7.
PROSITE; PS0138; BRC2_REPEAT; 6.
POLYmorphism; Repeat; 3D-structure.
REPEAT 981 1015
BRC2 1.
REPEAT 1192 1226
BRC2 2.
REPEAT 1394 1428
BRC2 3.
REPEAT 1491 1525
BRC2 4.
REPEAT 1623 1657
BRC2 5.
REPEAT 1924 1958
BRC2 6.
REPEAT 2004 2038
BRC2 7.
VARIANT 44 44
VARIANT 340 340
VARIANT 377 377
VARIANT 407 407
VARIANT 661 661
VARIANT 739 739
VARIANT 1038 1038
VARIANT 1198 1198
VARIANT 1257 1257
VARIANT 1392 1392
VARIANT 1520 1521
VARIANT 1583 1583
VARIANT 1613 1613
VARIANT 1686 1686
VARIANT 1799 1799
VARIANT 1881 1881
VARIANT 1894 1894
VARIANT 2141 2141
VARIANT 2392 2392
VARIANT 2605 2605
VARIANT 2648 2648
VARIANT 2717 2717
VARIANT 2729 2729
VARIANT 2814 2814
S -> F (IN STRAINS C57BL/6 AND 129/SV).
T -> P (IN STRAIN 129/SV).
N -> H (IN STRAIN C57BL/6).
H -> P (IN STRAIN C57BL/6).
I -> V (IN STRAIN C57BL/6).
P -> H (IN STRAIN C57BL/6).
I -> L (IN STRAINS C57BL/6 AND 129/SV).
QF -> RI (IN STRAIN C57BL/6).
Q -> P (IN STRAIN C57BL/6).
Q -> R (IN STRAIN C57BL/6).
FD -> CG (IN STRAIN C57BL/6).
R -> W (IN STRAIN C57BL/6).
C -> W (IN STRAIN C57BL/6).
C -> W (IN STRAIN C57BL/6).
S -> F (IN STRAIN C57BL/6).
S -> F (IN STRAIN 129/SV).
P -> L (IN STRAIN C57BL/6).
Q -> F (IN STRAIN 129/SV).
Q -> K (IN STRAIN C57BL/6).
S -> R (IN STRAIN C57BL/6).
K -> Q (IN STRAIN C57BL/6).
A -> P (IN STRAIN C57BL/6).
R -> C (IN STRAIN 129/SV).
L -> M (IN STRAIN 129/SV).
Q -> H (IN STRAIN C57BL/6).

Search completed: August 22, 2003, 18:56:42
Job time : 28 secs

[illegible]



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: August 22, 2003, 18:53:10 ; Search time 101 Seconds
(without alignments)
597.864 Million cell updates/sec

Title: US-10-030-613-1
Perfect score: 1254
Sequence: 1 MRSPGQGHVPEDGGLFLC.....SPKPKIOTRRGLPPSVSN 234

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_23.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriaph.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.		Query		Match Length DB ID		Description
		Score	%			
1	1241	99.0	243	4	Q8N116	Q8N116 homo sapien
2	1188	94.7	243	11	Q8CH26	Q8CH26 rattus norv
3	1184	94.4	243	11	Q8VIP4	Q8VIP4 mus musculu
4	903	72.0	197	11	Q8CH27	Q8CH27 rattus norv
5	891	71.1	192	11	Q8VIP5	Q8VIP5 mus musculu
6	634	50.6	199	11	Q8K4S2	Q8K4S2 rattus norv
7	622.5	49.6	198	11	Q91WQ4	Q91WQ4 mus musculu
8	146	11.6	315	3	Q9P5S9	Q9P5S9 neurospora
9	125.5	10.1	249	3	Q9P4A1	Q9P4A1 cryptococcu
10	98	7.8	504	4	Q8IUY9	Q8IUY9 homo sapien
11	96	7.7	277	3	O13703	O13703 schizosacch
12	96	7.7	455	11	Q920L8	Q920L8 mus musculu
13	96	7.7	535	11	Q9EQT7	Q9EQT7 mus musculu
14	94	7.5	746	4	Q9BU60	Q9BU60 homo sapien
15	94	7.5	813	16	Q9X2G3	Q9X2G3 thermotoga
16	94	7.5	1021	4	O15451	O15451 homo sapien

17	94	7.5	1052	4	Q96FT1	Q96FT1 homo sapien
18	94	7.5	1130	4	Q8I2L8	Q8I2L8 homo sapien
19	94	7.5	1284	4	O15450	O15450 homo sapien
20	93.5	7.5	688	11	Q8BIG8	Q8BIG8 mus musculu
21	93.5	7.5	732	11	Q8K2L2	Q8K2L2 mus musculu
22	93.5	7.5	751	16	Q8YUK6	Q8YUK6 anabaena sp
23	92	7.3	270	5	Q93194	Q93194 caenorhabdi
24	92	7.3	649	10	Q9M6E6	Q9M6E6 nicotiana t
25	90.5	7.2	220	10	Q9FGA3	Q9FGA3 arabidopsis
26	90.5	7.2	339	10	O8LJ34	O8LJ34 oryza sativ
27	90.5	7.2	351	10	Q9FGF8	Q9FGF8 arabidopsis
28	90	7.2	479	10	Q9M6E4	Q9M6E4 nicotiana t
29	90	7.2	1066	4	Q9H2N5	Q9H2N5 homo sapien
30	90	7.2	3310	4	Q9P222	Q9P222 homo sapien
31	90	7.2	4025	4	Q9NR13	Q9NR13 homo sapien
32	90	7.2	4911	4	O8NEZ4	O8NEZ4 homo sapien
33	88.5	7.1	501	5	Q9BKQ7	Q9BKQ7 caenorhabdi
34	88	7.0	498	10	Q9MAD6	Q9MAD6 arabidopsis
35	88	7.0	513	5	Q95Q27	Q95Q27 caenorhabdi
36	88	7.0	1015	10	Q9SRH8	Q9SRH8 arabidopsis
37	88	7.0	1090	10	O8LJ33	O8LJ33 oryza sativ
38	88	7.0	3329	11	Q8VHD0	Q8VHD0 mus musculu
39	87.5	7.0	1557	5	O96652	O96652 drosophila
40	87	6.9	671	10	Q9FXA2	Q9FXA2 arabidopsis
41	87	6.9	862	2	Q9EXA5	Q9EXA5 neisseria m
42	86.5	6.9	1112	2	Q8GD29	Q8GD29 bordetella
43	86	6.9	528	2	O9RLF1	O9RLF1 pleistomonas
44	85	6.8	243	12	O67788	O67788 human adeno
45	85	6.8	408	11	Q8BYE5	Q8BYE5 mus musculu

ALIGNMENTS

RESULT 1

Q8N116 PRELIMINARY; PRT; 243 AA.

ID Q8N116; DT 01-OCT-2002 (TREMBLrel. 22, Created)

AC Q8N116; DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DE Calciineurin inhibitor ZAKI-4 beta splice variant 1 (Calcineurin

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1] SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX PubMed=12102656;

RA Cao X., Kambe F., Miyazaki T., Sarkar D., Ohmori S., Seo H.;

RT "Novel human ZAKI-4 isoforms: Hormonal and tissue-specific regulation

RL Biochem. J. 0:0-0(2002)."

RN [2] SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Cao X., Kambe F., Miyazaki T., Ohmori S., Seo H.;

RT "Novel ZAKI-4 transcripts generated by alternative initiation and

RL splicing from a single gene on chromosome 6.";

DR EMBL; AY034085; AAK59805.1; to the EMBL/GenBank/DBJ databases.

DR EMBL; AY034086; AAK59806.1; "

SQ SEQUENCE 243 AA; 27247 MW; 780073AE29A973E4 CRC64;

Query Match 99.0%; Score 1241; DB 4; Length 243;
Best Local Similarity 99.1%; Pred. No. 5.8e-114;
Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSPGQGHVPEDGGLFLCCIDRDWAVTRCFEAFAQAITDFNDLPNSLFACNVHQSVF 60
|||||
Db 10 MRSPGQGHVPEDGGLFLCCIDRDWAVTRCFEAFAQAITDFNDLPNSLFACNVHQSVF 69

QY 61 EGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 120
 DB 70 EGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 129
 QY 121 YFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPVLYNDLLYAVAKLG 180
 DB 130 YFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPVLYNDLLYAVAKLG 189
 QY 181 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 234
 DB 190 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 243

RESULT 2

Q8CH26 PRELIMINARY; PRT; 243 AA.
 AC Q8CH26
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE ZAKI-4 beta.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RA Miyazaki T., Cao X., Kambe F., Ohmori S., Seo H.;
 RT "Cloning of rat ZAKI-4 beta cDNA."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF459023; AA015541.1; -
 SQ SEQUENCE 243 AA; 27311 MW; 00E05BDDICD0492D CRC64;

Query Match 94.7%; Score 1188; DB 11; Length 243;
 Best Local Similarity 94.4%; Pred. No. 9.5e-109;
 Matches 221; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFAEAEFAQAITDFNDLNSLFACNVHQSVE 60
 DB 10 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFAEAEFAQAITDFNDLNSLFACNVHQSVE 69
 QY 61 EGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 120
 DB 70 EGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 129
 QY 121 YFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPVLYNDLLYAVAKLG 180
 DB 130 YFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPVLYNDLLYAVAKLG 189
 QY 181 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 234
 DB 190 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 243

RESULT 3

Q8VIPA PRELIMINARY; PRT; 243 AA.
 AC Q8VIPA
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 23, Last annotation update)
 DE Calcineurin inhibitory protein ZAKI-4 beta.
 GN DSCRLL1 OR ZAKI-4 BETA.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.;
 RT "Calcineurin inhibitory protein ZAKI-4."

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB061525; BAB71956.1; -
 DR MGD: MGI:1858219; Dscr111.
 DR InterPro: IPR006931; Calciexpressin.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF04847; Calciexpressin; 1.
 DR Pfam: PF00076; rim; 1.
 SQ SEQUENCE 243 AA; 27332 MW; 2CD8C1A810291851 CRC64;

Query Match 94.4%; Score 1184; DB 11; Length 243;
 Best Local Similarity 94.0%; Pred. No. 2.3e-108;
 Matches 220; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFAEAEFAQAITDFNDLNSLFACNVHQSVE 60
 DB 10 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFAEAEFAQAITDFNDLNSLFACNVHQSVE 69
 QY 61 EGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 120
 DB 70 EGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 129
 QY 121 YFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPVLYNDLLYAVAKLG 180
 DB 130 YFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPVLYNDLLYAVAKLG 189
 QY 181 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 234
 DB 190 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 243

RESULT 4

Q8CH27 PRELIMINARY; PRT; 197 AA.
 AC Q8CH27
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE ZAKI-4 alpha.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RA Miyazaki T., Cao X., Kambe F., Ohmori S., Seo H.;
 RT "Cloning of rat ZAKI-4 alpha cDNA."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF459022; AA015540.1; -
 SQ SEQUENCE 197 AA; 21991 MW; D8C906B7A7F46F4B CRC64;

Query Match 72.0%; Score 903; DB 11; Length 197;
 Best Local Similarity 90.9%; Pred. No. 8e-81;
 Matches 170; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 48 NSLFACNVHQSVEGEESKEFGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIE 107
 DB 11 STLAVCVVDEVEFTNQEVKEKFEGLFRTYDECVTQLFKSFRRVRINFSNPKSAARARIE 70
 QY 108 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPV 167
 DB 71 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQPAKQFLISPPSSPPVQWOPINDATPV 130
 QY 168 LNYDLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPG 227
 DB 131 LNYDLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPG 190
 QY 228 LPPSVSN 234
 DB 191 LPPSVSN 197

RESULT 5

Db 5 DFNFNFSLLIACVANGDVFSESETRAKFESLFRTYDKDITFQYFKSFRVRINFSNPLSA 64

Qy 102 ARARIELHETQFRGKKLKLFAQVOTPEYDGDGLHLAPQAKQFLISPPSPPPVGWQPI 161

Db 65 ADARLRHKHTEFLGKEMKLYFAOTLHIGSS----HLAPPNDPKQFLISPPASPPVGWKQV 120

Qy 162 NDATPVNLDLYAVAKLGPGEKGYELHAGTETSPSVVHVCDSDTEEEEDPKTS-----215

Db 121 EDATPVNLDLYAISKLGPGKEKGYELHAATDTTPSVVHVCDSDTEEEEDPKTS-----215

Qy 216 PKPKIIQTRRPGPLP 230

Db 181 PKPKIIQTRRPEYTP 195

RESULT 7

Q91WQ4 ID Q91WQ4 PRELIMINARY; PRT; 198 AA.

AC Q91WQ4;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Down syndrome critical region homolog 1 (Human).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Strausberg R.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC013551; AAH13551.1; -.

DR InterPro: IPR006931; Calciopressin.

DR Pfam: PF04847; Calciopressin; 1.

SQ SEQUENCE 198 AA; 22851 MW; F017C68F18AC187 CRC64;

Query Match 49.6%; Score 622.5; DB 11; Length 198;

Best Local Similarity 62.9%; Pred. No. 3.2e-53;

Matches 122; Conservative 23; Mismatches 40; Indels 9; Gaps 2;

Qy 42 DFDLPLNSLFPACNVHOSVTEGESKKEFGLFRTYDCVTFQLFKSFRVRINFSNPKSA 101

Db 5 DFNFNFSLLIACVANGDVFSESETRAKFESLFRTYDKDITFQYFKSFRVRINFSNPLSA 64

Qy 102 ARARIELHETQFRGKKLKLFAQVOTPEYDGDGLHLAPQAKQFLISPPSPPPVGWQPI 161

Db 65 ADARLRHKHTEFLGKEMKLYFAOTLHIGSS----HLAPPNDPKQFLISPPASPPVGWKQV 120

Qy 162 NDATPVNLDLYAVAKLGPGEKGYELHAGTETSPSVVHVCDSDTEEEEDPKTS-----TSP 216

Db 121 EDATPVNLDLYAISKLGPGKEKGYELHAATDTTPSVVHVCDSDTEEEEDPKTS-----TSP 216

Qy 217 PKPKIIQTRRPGPLP 230

Db 181 PKPKIIQTRRPEYTP 194

RESULT 8

Q9P5S9 ID Q9P5S9 PRELIMINARY; PRT; 315 AA.

AC Q9P5S9;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Nebula related protein.

DE B5032.230.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariidae; Sordariaceae; Neurospora.

OX NCBI_TaxID=5141;

RN [1]

RP SEQUENCE FROM N.A.

Tue Aug 26 09:11:04 2003

```

RA Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL EMBL; AL355932; CAB91442.1; -
DR InterPro; IPR006931; Calcipressin.
DR Pfam; PF04847; Calcipressin; 1.
SQ SEQUENCE 315 AA; 34194 MW; 76792987653C3AB7 CRC64;

Query Match 11.6%; Score 146; DB 3; Length 315;
Best Local Similarity 31.0%; Pred. No. 4.6e-06;
Matches 49; Conservative 21; Mismatches 60; Indels 28; Gaps 7;

QY 42 DFNDLP-----NSIFACNVHQ-SVFGEESKEKEFGLFRYYDCV--TFOLF--- 85
DB 48 DLSNLPPLSQPTPPSNTLIFTNINSLDVFSADN-----LQTRIDLSQTAFIAWSP 99
QY 86 -KSFRRVRINFSPKSAARARLHETQFRGKCLKLYFAQVQTPETDGDGKLHLAPPQPAK 144
DB 100 LKSFRRIVTFDEQAATIAVRVWDGEAILGRCRVFGQTPIDVSAADKHLALPDAGK 159
QY 145 QELISPPSPPPVGV-OPINDA--TPVLYNDLLYAVAKL 179
DB 160 LFFISPPSPPPHDEQRMEDAPNTWVHAEALAEALAKL 197

RESULT 9
Q9P4A1 PRELIMINARY; PRT; 249 AA.
AC Q9P4A1
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN CBP1.
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H99;
RA MEDLINE=20359261; PubMed=10899116;
RX Gorlach J., Fox D.S., Cutler N.S., Cox G.W., Perfect J.R., Heitman J.;
RT "Identification and characterization of a highly conserved calcineurin-
binding protein, CBP1/calcipressin, in Cryptococcus neoformans.";
RL EMBL J. 19; 3618-3629(2000).
DR EMBL; AF230799; AAF63734.1; -
DR InterPro; IPR006931; Calcipressin.
DR Pfam; PF04847; Calcipressin; 1.
SQ SEQUENCE 249 AA; 27293 MW; 78757183FA5EBF09 CRC64;

Query Match 10.1%; Score 126.5; DB 3; Length 249;
Best Local Similarity 38.0%; Pred. No. 0.00028;
Matches 30; Conservative 12; Mismatches 28; Indels 9; Gaps 2;

QY 118 LKLYFAQVQTPETDGDGKLHLAPPQPAKOFILSPSPSPVGPQINDATP-----VLNYDLY 174
DB 107 LRLHYLPPTLPNDPATTHLAPPPLPHNLISPPGSPGEGWEPAAEAPNRILLPEDLQ 166
QY 175 AVAKLGPGEKYLHAGTES 193
DB 167 AL-----ETLENSGSKA 179

RESULT 10
Q8IU9 PRELIMINARY; PRT; 504 AA.
ID Q8IU9
AC Q8IU9
DT 01-MAR-2003 (Tremblrel. 23, Created)

```

```

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to yes-associated protein, 65 kDa.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Pancreas;
RC Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038235; AAH38235.1; -. 6145F7049ED338AE CRC64;
SQ SEQUENCE 504 AA; 54461 MW; 54461 MW; 54461 MW; 54461 MW; 54461 MW;

Query Match 7.8%; Score 98; DB 4; Length 504;
Best Local Similarity 31.1%; Pred. No. 0.45;
Matches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;

QY 139 PPQPAKQFLISPPSPVGPQINDATPVLNYDLYAVAKLGPGEKYLHAGTESPPSV- 197
DB 8 PPQPAQOGQOPSPQPGQGPSP-----GPGQ--PAPAATQAAP 48
QY 198 -----VVHV-CDSDIEEE-----DPKTSPPKPIIOTRRPGLPPS 231
DB 49 PAGQIVHVRGDSSETDLEALFNAMNPKTANVPQTPMRLKLPDS 94

RESULT 11
Q13703 PRELIMINARY; PRT; 277 AA.
ID Q13703
AC Q13703
DT 01-JAN-1999 (Tremblrel. 09, Created)
DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN SPAC13F5.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Brown D., Churcher C.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99091; CAB11767.1; -.
DR GeneDB SPombe; SPAC13F5.04c; -.
RW Hypothetical protein.
SQ SEQUENCE 277 AA; 29622 MW; 540CF08359A06546 CRC64;

Query Match 7.7%; Score 96; DB 3; Length 277;
Best Local Similarity 22.4%; Pred. No. 0.32;
Matches 46; Conservative 22; Mismatches 77; Indels 60; Gaps 7;

QY 79 CVTFQFLKSPRRVRINF-----SNPKSAAR-----ARIELHETQFRGK 116
DB 2 CFFELVSSFAILRISFLSITKICKCKLRALQVANPQKAKTPSNHATELQOOSTNST 61
QY 117 KKLKLYFAQVQTPETDGDGKLHLAPP--QPAKOFILSPSPSPVGPQ-----INDA 164
DB 62 TLPTQEAIVETNASASHETSFALPTTSPAASLSLSPKSAVSEPNAVADVKSLSSTPA 121
QY 165 TPVLYNDLLYAVAKLGPGEKYE-----LHAGTESPSV-----VVHVCDSDIEE 208
DB 122 APQLN-----SPSHSYEPTTFPTTSTITENLPTIDPTRTRSSSHIQSLSPESKQ 171

```

```

DR MGD; MGI:1927803; Fcamr.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000508; SigPase.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00409; Ig_1.
DR PROSITE; PS00501; SPASE_I_1; 1.
KW Receptor.
SQ SEQUENCE      535 AA; 57696 MW; B275B77C70151C75 CRC64;

Query Match
Best Local Similarity   7.7%; Score 96; DB 11; Length 535;
Matches    29; Conservative    19; Mismatches    50; Indels    18; Gaps    4.

QY 124 QVOTPTDGDK-----LHLAPPAKQFLLSPSPS-----SPPGVQPIINDATPVNLNLLYAV 176
Db 323 OGTTTPTDGRDIDVRVSPEAPRKTGTTTRSALISBEHVWETLODKTEYSKQOMLHSL 382
QY 177 AKLGPEKYE-----LIAGTESTSVVVHVCDSDIEEEDPKTSKPKKIIQTRRP 226
Db 383 EELSPPASQAOTLNATCLEVASEGRSI-----DGSLENTTBSSPPTPSQLSVAGP 433

RESULT 14
Q9BU60 PRELIMINARY; PRT; 746 AA.
AC Q9BU60;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Straussberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC02875; AA02875.1; -.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE      746 AA; 79187 MW; 35A186C5A7ADEA8B CRC64;

Query Match
Best Local Similarity   7.5%; Score 94; DB 4; Length 746;
Matches    41; Conservative    20; Mismatches    80; Indels    32; Gaps    6;

QY 61 EGESKEKFGLFRFYDDCVTFOLKFSFRVRINFSNPKSAAARIELHETQPRGKKKL 120
Db 539 EEEEEEEFEFEFEFG-----ELEEEEEEEEEELEEVE----- 577
QY 121 YFAQVTPETDGNLHAPPQAKQFLIIPSPSPVPVGMQPIINDATPVNLNLLYAVAKLG 180
Db 578 ---DLFFGTAGGEVEEGCAPPTLPALPPESP-KVQPEPEPEP---GLLEVEE-- 627
QY 181 PGKEYELHAGTESTPSPVVHVCDSDIEE-EDPKTSKPKKIIQTRRGPPSPV 232
Db 628 PGTEERGAADTAPTLAEPALPSQGEVEREGESPAAGPPPQELVEEPSAPTLL 680

RESULT 15
Q9X2G3 PRELIMINARY; PRT; 813 AA.
AC Q9X2G3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cellobiose-phosphorylase.
GN TM1848.
```

Tue Aug 26 09:11:04 2003

```
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
  Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
  McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
  Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
  Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
  Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
  genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001822; AAC36910.1; -.
DR TIGR; TM1848; -.
DR InterPro; IPR000169; SHprot_acsite.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Complete proteome.
SQ SEQUENCE 813 AA; 93499 MW; 55FC07E7329D8C41 CRC64;

Query Match          7.5%; Score 94; DB 16; Length 813;
Best Local Similarity 24.5%; Pred. NO. 2.1; Indels 34; Gaps 9;
Matches 45; Conservative 27; Mismatches 78;

QY 51  FACNVHQSVEGESKEFGLFRTYDCCVTQOL--FKSFRVRVINFSPKSAARARIEL 108
DB 126 FTGEVHYLVLENKAEPKIKLFSFIEFLNALLDDMTNFOR---NYSTGEVEIEGVIY 182
QY 109 HETQFRGKKLKLFPQVQTP-----ETDGDK-LHLAPPQPAKOFLL--SPSSPPVGMQPI 161
DB 183 HKTEYRRNRHYAFYSVNQPIDGFDYDRESFIGLYSGFEAPQAVVEGKPRNSVASCWAPI 242
QY 162 NDATPVLNYDLLYAVAKLGPGEKYEHLHAGTESTPSVVVHVHVCDSIDIEEDPKTSPKPII 221
DB 243 ASH-----YLEIELAPSEKKEL-----IFIL-GYVENPEEEKWE-KPGVI 280
QY 222 QTRR 225
DB 281 NKKR 284
```

Search completed: August 22, 2003, 18:58:32
Job time : 106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 07:02:29 ; Search time 376 Seconds
(without alignments)
6705.519 Million cell updates/sec

Title: US-10-030-613-3
Perfect score: 934
Sequence: 1 agaatcatacttcacggaa.....gtataaccattatttagaa 934

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2004.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	934	100.0	934	22	AAF25338
2	934	100.0	3159	24	ABAG1457
3	720.8	77.2	3184	24	AAD30157
4	720.8	77.2	3184	24	ABL61768
5	720.8	77.2	3184	24	ABL61768
6	422	45.2	1021	24	ABAG1458
7	413.2	44.2	594	24	AAD30154
8	279.2	29.9	828	24	AAD30158

9	257.6	27.6	615	24	ABA91463
10	231.6	24.8	720	24	AAD30159
11	211	22.6	597	24	AAD30152
12	209.6	22.4	597	24	AAD30153
13	197.6	21.2	599	24	AAD30151
14	194.4	20.8	2212	24	AAD30156
15	187	20.0	2331	24	AAD30155
16	187	20.0	2346	24	ABT11031
17	187	20.0	2355	25	ABX63255
18	184.2	19.7	2348	24	AAS94805
19	181.6	19.4	2358	21	AAF18328
20	177.4	19.0	2173	24	ABN96843
21	177.4	19.0	2174	20	AAO1282
22	167	17.9	363	22	AAI88711
23	133	14.2	227	21	AAI11239
24	126.8	13.6	531	24	ABA91461
25	116	12.4	1820	23	ABL29301
26	116	12.4	12550	23	ABL29300
27	93.2	10.0	442	21	AAC01774
28	83.6	9.0	365	24	ABS69673
29	83.6	9.0	412	22	ABA56172
30	83.6	9.0	412	22	AAK04364
31	83.6	9.0	412	22	AAI14449
32	83.6	9.0	412	22	AAI35821
33	83.6	9.0	412	22	AAI04272
34	83.6	9.0	412	23	ABS29505
35	83.6	9.0	412	24	ABS04419
36	83.6	9.0	446	22	ABA3406
37	83.6	9.0	446	22	ABA53852
38	83.6	9.0	446	22	ABA23598
39	83.6	9.0	446	22	AAK02113
40	83.6	9.0	446	22	AAK27563
41	83.6	9.0	446	22	AAI12148
42	83.6	9.0	446	22	AAI33493
43	83.6	9.0	446	22	AAI02061
44	83.6	9.0	446	23	ABS27133
45	83.6	9.0	446	24	ABS02041

ALIGNMENTS

RESULT 1
AAF25338
ID AAF25338 standard; cDNA; 934 BP.
XX
AC AAF25338;
XX

DT 30-APR-2001 (first entry)

XX Nucleotide sequence of a human detoxification protein.

Human; detoxification protein; DETX; cancer; leukaemia; melanoma;
adenocarcinoma; autoimmune disorder; inflammatory disorder;
rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
psoriasis; ulcerative colitis; infection; cell proliferative disorder;
actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 20..724

XX sig_peptide /product= "detoxification protein"
XX FT /*tag= a
XX FT 20..100
XX FT /*tag= b

XX WO200104305-A2.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000WO-US18509.

XX

Applied

07-JUL-1999; 99US-0142678.
(INCY-) INCYTE GENOMICS INC.
Tang YT, Yue H;
WPI: 2001-147193/15.
P-PSDB; AAB31788.
New human detoxification protein and polynucleotide, useful for diagnosis, prevention and treatment of autoimmune/inflammatory disorders and cell proliferative disorders including cancer -
Claim 5; Page 79; 79pp; English.
The present sequence encodes a human detoxification polypeptide (DETX). DETX and its (ant)agonists are useful for preventing or treating disorders associated with decreased or increased expression or activity of DETX. DETX polypeptides are useful for screening compounds that specifically binds to DETX and for identifying (ant)agonists. Diseases prevented, treated and diagnosed include cancers (e.g. leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow, breast, kidney, liver, pancreas, prostate and uterus), autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma, atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative colitis), bacterial, fungal, parasitic infections and cell proliferative disorders (e.g. actinic keratosis, arteriosclerosis, cirrhosis and hepatitis). Anti-DETX antibodies may be used as antagonists, as a targeting or delivery mechanism for bringing pharmaceutical agents into contact with cells or tissues expressing DETX and for diagnosis of DETX-related disorders.
Sequence 934 BP: 247 A; 244 C; 206 G; 237 T; 0 other;
Query Match 100.0%; Score 934; DB 22; Length 934;
Best Local Similarity 100.0%; Pred. No. 4.4e-294;
Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AGAATCATACTTCATCGGAATGAGGAGCCAGGACGAGGACACGTCCTCGAAGATGG 60
1 AGAATCATACTTCATCGGAATGAGGAGCCAGGACGAGGACGAGGACGTCCTCGAAGATGG 60
61 AGGACTTTTCTTACTGTGCTGCATAGACAGGAGCTGGGCTGTCACGTGTTTTCGAGA 120
61 AGGACTTTTCTTACTGTGCTGCATAGACAGGAGCTGGGCTGTCACGTGTTTTCGAGA 120
121 AGAAGCCTTTCACGCAATCACTGACTTCAATGACCTCCCAACTCGTGTGTTGCGTGCAA 180
121 AGAAGCCTTTCACGCAATCACTGACTTCAATGACCTCCCAACTCGTGTGTTGCGTGCAA 180
181 TGTTCCACCACTGTCAGTGTGTTGAAGGAGAGAGAGCAAGAAATTTGAGGGACTGTTTCG 240
181 TGTTCCACCACTGTCAGTGTGTTGAAGGAGAGAGAGCAAGAAATTTGAGGGACTGTTTCG 240
241 GACTTATGATGACTGTGAGCTGCCAGCTATTTAAGAGTTTCAGAGCTGTCCTGATATAA 300
241 GACTTATGATGACTGTGAGCTGCCAGCTATTTAAGAGTTTCAGAGCTGTCCTGATATAA 300
301 CTTTCAGCAATCCTAAATCTGCACCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAG 360
301 CTTTCAGCAATCCTAAATCTGCACCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAG 360
361 AGGGAAAAAATTTAAGCTCTACTTTGACAGCTTCAGACTCCAGACAGACAGATGGAGACAA 420
361 AGGGAAAAAATTTAAGCTCTACTTTGACAGCTTCAGACTCCAGACAGACAGATGGAGACAA 420
421 ACTGCACTTGGGTGCCACCCAGCCTGCCAAACAGTTTCTCATCTCGCCCCCTTCCTCCC 480
421 ACTGCACTTGGGTGCCACCCAGCCTGCCAAACAGTTTCTCATCTCGCCCCCTTCCTCCC 480
481 ACTGTTGGCTGCGACCCCATCAACGATGCCAGCCAGTCCTCAACTATGACCTCTCTCTA 540
481 ACTGTTGGCTGCGACCCCATCAACGATGCCAGCCAGTCCTCAACTATGACCTCTCTCTA 540


```

Db 763 ATTTGAGGACTGTCGGACCTATGACGAATGTGTGAGCTTTGAGCTGTTTAAAGAGTTT 704
QY 283 CAGACGTG-TCCGTATAAATTCAGCAATCTTAAATCTGCAGCCG--AGCTAGATAGA 339
Db 703 CCGACGGTTCGGGAATAAATTCAGCACCCCAAGCTGCAGCCCGTGGCCGGATAGA 644
QY 340 GCTTCA--TGAACCAATTCAGAGGAGGAAATTAAGAGCTCTACTTTGCACAGTTTCAG 397
Db 643 GCTTATGTGGACCCAGTTCATGAGGAGGAGAGTGTGAAGCTCTACTTTGCACAGTTCAG 584
QY 398 ACTCCAGACAGATGG-AGACAACTGCA-CTTGGCTTCCACCCAGCTGCCAAACAGT 455
Db 583 ACCCCAGACAGATGGTAGAGAGTGTGAAGCTCTACTTTGCACAGTTCAG 455
QY 456 TTCTCATCTCGCCCTCTCTCCACCTGTGCTGGCAGCCCATCAACGATGCCACGC 524
Db 523 TCCTCATCTACCCCTCTCATCTCTCCGTTGGTGAAGCTTATCAGGATGCCACAC 464
QY 516 CAGTCTCAAC-TATGACCTCTCTATGCTGTGCGCAAACTAGGACAGGAGAGATAT 574
Db 463 CAGTCTCAACGTACGACCTCTCTTATGCGGTGCGCAAACTAGGACAGGAGAGATAT 404
QY 575 GAGTCCATCAGGAGTGTGCTCCACCCAGTGTGCTGTCGACGTGTGCGACAGTGAC 634
Db 403 GAGTGTGATCGGGAAGTGTGCTACCCAGGCTGTGTCGACGTGTGACAGCGAC 344
QY 635 ATAGAGGAAGAGAGGACCAAGACTTCCCAAGGCAAAATCATCAAAATTCGCTTAAAT-TCC 813
Db 343 TTGGAGGAGAGGAGATCCAAAGACTTCCCAAGGCAAAATCATCAAAATTCGCTTAAAT-TCC 813
QY 695 CTTGGGCTGCCACCCCTCGTGTGCTCAACTGAGTGTGCTCTCTTCGATATATAGCCGTC 754
Db 283 CTTGGGCTGCCACCCCTCGTGTGCTCAACTGAGTGTGCTCTCTTCGATATATAGCCGTC 754
QY 755 TCCTCTTTATGCTGTTTCCCTCTGTTGTTGTCGCAAAATTCGCTTAAAT-TCC 813
Db 230 TCCTC-TTATCATGCTTTTCCCTCTGTTGTCGCAAAATTCGCTTAAAT-TCC 813
QY 814 TGGGTGTTGGTGTGTTGAGATTCCTCTCTGTTA-TCAAGCCTCTCGGCAAAAGGCT 872
Db 171 TGGGTATTTGGTGTGTTGAGATTCCTCTCTGTTA-TCAAGCCTCTCGGCAAAAGGCT 872
QY 873 AGGAAAGTGATATGCTCTGATCATATATATATATATATATATATATATATATATAT 932
Db 111 AGGAAAGTGATATGCTCTGATCATATATATATATATATATATATATATATATAT 932
QY 933 AA 934
Db 51 AA 50

```

```

RESULT 7
AAD30154
ID AAD30154 standard; DNA; 594 BP.
AC AAD30154;
XX AAD30154;
DT 17-MAY-2002 (first entry)
DE Mouse MCIP associated DNA #4.
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; mouse; gene; ds.
OS Mus musculus.
XX Key
XX Location/Qualifiers
XX 1..594
XX /*tag= a
XX /product= "Mouse MCIP associated protein #4"
XX
XX WO200204491-A2.

```

```

XX PD 17-JAN-2002.
XX PF 06-JUL-2001; 2001WO-US21662.
XX PR 07-JUL-2000; 2000US-216601P.
XX PR 13-FEB-2001; 2001US-0782953.
XX PA (TEXA) UNIV TEXAS SYSTEM.
XX PA (WILL) WILLIAMS S R.
XX PA (ROTH) ROTHERMEL B.
XX PI Williams SR, Rothermel B.
XX WPI; 2002-179698/23.
XX P-PSDB; AAE18913.
XX
XX Screening for modulators of muscle calcineurin interacting protein
XX (MCIP) binding, expression or phosphorylation, useful for treating
XX cardiac hypertrophy or heart failure, comprises mixing MCIP,
XX calcineurin and a test compound.
XX
XX Disclosure: Page 153-154; 174pp; English.
XX
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX complex with the catalytic subunit of calcineurin and increased levels
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX transcription of certain target genes. The invention also relates to
XX methods for identifying modulators of MCIP binding, expression or
XX phosphorylation. Inhibitors or promoters of MCIP binding, expression or
XX may be used for treating cardiac hypertrophy and heart failure.
XX Antibodies to MCIP can be used in characterizing the MCIP content of
XX healthy and diseased tissues and subsequently for determining the
XX presence or absence of cardiomyopathy or as predictor of heart disease.
XX The present sequence is mouse MCIP associated DNA.
XX Note: This sequence has been described as mouse MCIP1 in the
XX specification, however the sequence seems to be a polynucleotide
XX encoding a MCIP associated protein.
XX
XX Sequence 594 BP; 148 A; 165 C; 149 G; 132 T; 0 other;

```

```

Query Match 44.2%; Score 413.2; DB 24; Length 594;
Best Local Similarity 83.5%; Pred. No. 4.5e-124;
Matches 469; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 163 CTCGTGTTGGCTGCAATGTTCCACGACGAGTGTGTTGAAGGAGAGAGAGAGAGAG 222
Db 33 CACTCTGGTCGCTGTGTGTGATGTGGAGGTCTTTACCAATCAGGAGGTAAAGGAAA 92
QY 223 ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTTAAGAGTTT 282
Db 93 ATTCGAGGAGCTGTTCCGAGCCTATGATGAATGTGTGAGCTTCCAGCTGTTTAAAGTTT 152
QY 283 CAGAGCTGTCCGTATAAATTCAGCAATCTTCAAGCAATCTCCAGCCAGTATAGATAGCT 342
Db 153 CCGACGGTTCGAATAAATTTCCAGCCATCCCAATCTCGACCCGTCGCGGATAGAGCT 212
QY 343 TCATGAAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 402
Db 213 TCATGAGACTCAGTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 272
QY 403 AGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 462
Db 273 AGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 332
QY 463 CTGCCCCCTCTCTCCACCTGTGTGGCTGGCAGCCCATCAAGATGCCAGCCAGCTCCT 522
Db 333 CTACCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 392
QY 523 CAACTATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 582
Db 393 CAACTATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 452

```


Query Match 21.2%; Score 197.6; DB 24; Length 599;
 Best Local Similarity 64.6%; Pred. No. 1.2e-53;
 Matches 332; Conservative 0; Mismatches 169; Indels 13; Gaps 2;

QY 138 TCACTGACTTCAATGACCTCCCAACTCGTGTGGTGGTGAATGTCACCAAGTCAGTGT 197
 DB 87 TTAGGACITTAGCTACAAATTTAGCTCCCTGATTTGTTGGTGAACGATGATGCT 146
 QY 198 TTGAAG 257
 DB 147 TCAGCGAAAGTGAAG 206
 QY 258 TGACGTTCCAGCTATTAAAGAGTTTCAGAGCTGTCCGTATAAAGTTCAGCAATCCTAAT 317
 DB 207 CCACCTCCAGTATTAAAGAGCTTCAACAGCTGTCCGGATAAAGTTCAGCAACCCCTTAT 266
 QY 318 CTGAGCCCGAGCTAGATAGAGCTTCATGAACCAATTCAGAGGAGAGAGAGAGAGAG 377
 DB 267 CTGAGCCGATGCCAGCTCGGCTGCACAGACCGAGTTCCTGGGAGAGAGATGAAGT 326
 QY 378 TCTACTTTGACAGAGTTTCAGACTCCAGAGACAGATGGAGACAACTGCACCTGGCTCCAC 437
 DB 327 TGTATTTGCTCAGACT-----TTACACATAGGAAGTTTCACACCTGGCTCC-G 373
 QY 438 CCCAGCTGCCAAACAGTTCTCATCTCGCCCTTCTCCACCTGTGGTGGTGGCAGC 497
 DB 374 CCAATCCCGACAAACAGTCTCTCATCTCCCTCCGCTCTCTCCGCTGGTGGTGGNAAC 433
 QY 498 CCATCAACGATGCCAGCGAGTCTCAACTATGATGACCTCTCTATGCTGTGGCCAACTAG 557
 DB 434 AAGTAGAGATGCCACCCCGCTCATAAATTACGATCTTTATATGCTCATCTCCAAAGCTGG 493
 QY 558 GACCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
 DB 494 GCGCAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
 QY 618 ACGTGTGCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
 DB 554 ACGTGTGCGAGAGTGACCAAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 587

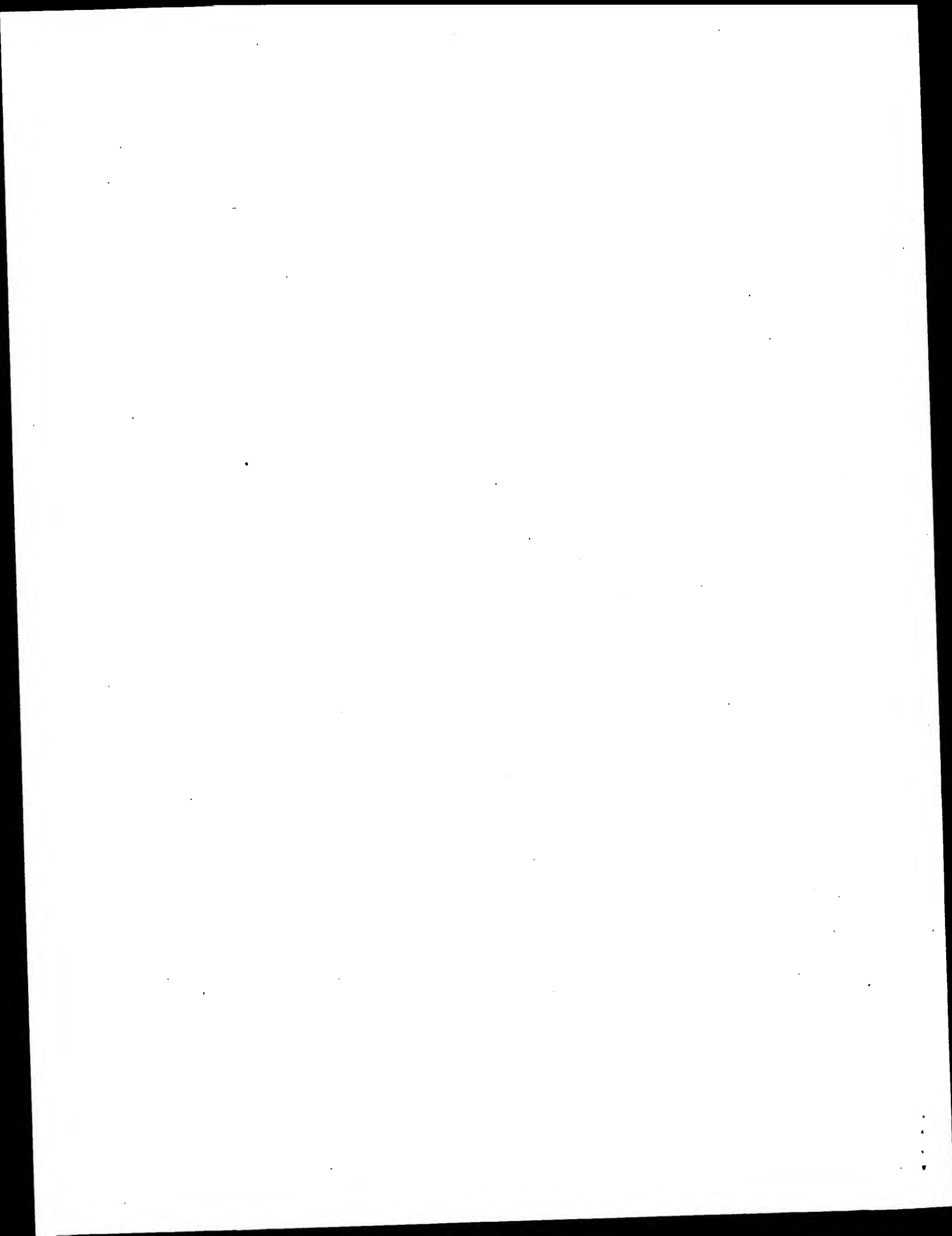
RESULT 14
 AAD30156
 ID AAD30156 standard; DNA; 2212 BP.
 AC AAD30156;
 XX 17-MAY-2002 (first entry)
 DT Human MCIP associated DNA #2.
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 25..618
 FT /tag= a
 FT /product= "Human MCIP associated protein #2"

WO200204491-A2.
 17-JAN-2002.
 06-JUL-2001; 2001WO-US21662.
 07-JUL-2000; 2000US-216601P.
 13-FEB-2001; 2001US-0782953.
 (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;
 XX WPI; 2002-179698/23.
 DR P-PSDB; AAE18915.
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 PS Example 1; Page 159-161; 174pp; English.
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 CC Note: This sequence has been described as murine MCIP splice variant in
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.
 XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;

Query Match 20.8%; Score 194.4; DB 24; Length 2212;
 Best Local Similarity 60.8%; Pred. No. 2.9e-52;
 Matches 365; Conservative 0; Mismatches 211; Indels 24; Gaps 2;

QY 143 GACTTCAATGACCTCCCAACTCGTGTGGTGGTGAATGTCACCAAGTCAGTGTGA 202
 DB 37 GACTGACGAGCTGCCAGCGCCACCATGCTGTCACCTGGACCGCGGTGTCGTG 96
 QY 203 GGAG 262
 DB 97 GACGCGCTGTGCGCGGCAAAATTTAGTTCCTCTTTAGGACGTATGACAGGACATCACC 156
 QY 263 TTCCAGCTATTAAAGAGTTTCAGACGTGTCCGTATAAATTCAGCAATCTTAATCTGCA 322
 DB 157 TTTCAGTATTATAGAGCTTCAACAGAGTCAAGATAAATTCAGCAACCCCTTCTCCGCA 216
 QY 323 GCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGAGAGAGAGAGAGAGAGAG 382
 DB 217 GCAGATGCCAGGCTCCAGCTGCATAAGAGTGTGTTCTGGAGAGAGAGAGAGAGAGAG 276
 QY 383 TTTCACAGGTTTCAGACTCCAGAGACAGATGGAGACAACTGCACTTGGCTCCACCCAG 442
 DB 277 TTGCTCAG-----ACCTTACACATAGGAAGTTCACACCTGGCTCCGCCAAAT 324
 QY 443 CTGCGCAACAGTTTCTCATCTCGCCCTTCTCCACCTGTGGCTGGCAGCCCATC 502
 DB 325 CCAGACAAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGGAACAGTG 384
 QY 503 AACGATGCCAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCAAACTAGGACCA 562
 DB 385 GAAGATGCGACCCAGTCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 444
 QY 563 GGAGAGAGATGAGTCCATGCGAGGAGTCCACCCCAAGTGTGCTGTCACCTG 622
 DB 445 GGGGAAAGATGATGAATTCAGCGCAGCTGCACACCTCCACAGGCTGGTGTCCATGTA 504
 QY 623 TCGGACAGTGACATAGAG-----GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
 DB 505 TGTGAGAGTGAATCAAG 564
 QY 671 CCAAAATCATCCAAACTCGGCTGCTGGCCTGCCACCTCCGCTGCTCCAACTGAGCTGCC 730



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 11:27:06 ; Search time 322 Seconds
(without alignments)
6521.071 Million cell updates/sec

Title: US-10-030-613-3

Perfect score: 934

Sequence: 1 agaatcatacttcaggaa.....gtataaccattatttagaa 934

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 151743 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	720.8	77.2	3184	10	US-09-954-456-497
2	720.8	77.2	3184	10	US-09-782-953-17
3	413.2	44.2	594	10	US-09-782-953-8
4	279.2	29.9	828	10	US-09-782-953-20
5	231.6	24.8	720	10	US-09-782-953-23
6	211	22.6	597	10	US-09-782-953-2
7	209.6	22.4	597	10	US-09-782-953-5
8	197.6	21.2	599	10	US-09-782-953-1
9	194.4	20.8	2212	10	US-09-782-953-14
10	187	20.0	2331	10	US-09-782-953-11
11	187	20.0	2355	13	US-10-044-090-255
12	187	20.0	2355	14	US-10-084-817-36
13	181.6	19.4	2358	9	US-09-925-302-347
14	177.4	19.0	2173	10	US-09-880-107-3340
15	177.4	19.0	2173	14	US-10-102-524-1746
16	177.4	19.0	2409	14	US-10-198-846-13070

Sequence 736, App
Sequence 10389, A
Sequence 2064, Ap
Sequence 1126, Ap
Sequence 114005
Sequence 27019, A
Sequence 809, App
Sequence 17592, A
Sequence 18808, A
Sequence 11, Appl
Sequence 6418, Ap
Sequence 708, App
Sequence 346, App
Sequence 346, App
Sequence 154378
Sequence 154379
Sequence 225823
Sequence 225824
Sequence 211428
Sequence 781, App
Sequence 191694
Sequence 3, Appli
Sequence 115812
Sequence 358, App
Sequence 358, App
Sequence 358, App
Sequence 358, App
Sequence 7591, Ap
Sequence 37189, A

17 83.6 9.0 365 10 US-09-728-445-736
c 18 83.6 9.0 412 9 US-09-864-761-10388
c 19 83.6 9.0 446 9 US-09-864-761-2064
c 20 79.8 8.5 518 14 US-10-102-524-1126
c 21 64 6.9 3404 13 US-10-027-632-114005
c 22 56 6.0 111 9 US-09-864-761-27019
c 23 47.2 5.1 486 9 US-09-864-761-17592
c 24 46.4 5.0 123 9 US-09-864-761-18808
c 25 43.6 4.7 85 9 US-09-920-671-11
c 26 37.6 4.0 139257 11 US-09-764-891-6418
c 27 37.2 4.0 17904 14 US-10-205-428-708
c 28 37.2 3.8 671 14 US-10-184-644-346
c 29 35.4 3.8 671 14 US-10-184-644-346
c 30 35.4 3.8 873 13 US-10-027-632-154378
c 31 35.2 3.8 873 13 US-10-027-632-154379
c 32 35.2 3.8 589 13 US-10-027-632-225823
c 33 34.8 3.7 589 13 US-10-027-632-225824
c 34 34.8 3.7 590 13 US-10-027-632-211428
c 35 34.4 3.7 7972 12 US-10-017-161-781
c 36 34.2 3.7 533 13 US-10-027-632-191694
c 37 34 3.6 378361 11 US-09-901-136-3
c 38 34 3.6 621 13 US-10-027-632-115812
c 39 33.4 3.6 1049 12 US-10-140-472-358
c 40 33.4 3.6 1049 12 US-10-141-761-358
c 41 33.4 3.6 1049 14 US-10-123-155-358
c 42 33.4 3.6 1049 15 US-10-146-731-358
c 43 33.4 3.6 1049 15 US-09-918-995-7591
c 44 33.2 3.6 361 11 US-09-918-995-7591
c 45 33.2 3.6 432 11 US-09-918-995-37189

ALIGNMENTS

RESULT 1

US-09-954-456-497
; Sequence 497, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 869290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR FILING DATE: 2000-09-18
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 497
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-497

```

; SEQ ID NO 17
;
; LENGTH: 3184
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (205)..(780)
;
; 785-09-782-953-17

```

Query Match	77.2%	Score 720.8	DB 10	Length 3184
Best Local Similarity	95.9%	Pred. No. 1.4e-246	Indels 0	Gaps 0
Matches 740	Conservative	0	Mismatches 32	
163	CTCGTTGTTTCGGTGCATGTTCCACGATGTCACGATGTCAGTCTGTTTGAAGGAGAAGAGAGACCAAGAAAA	222		
222	CACCTCGTGGTGGCTGTGTGGTGGATGTCGAGGTCTTTACCAATCAGGAGGTTAAGGAAAA	281		
223	ATTTGAGGAGCTGTTTTCGGACTTATGATGACTGTGTGACCTTCCAGCTATTTAAGAGTTT	282		
282	ATTTGGGGAGCTGTTTTCGGACTTATGATGACTGTGTGACCTTCCAGCTATTTAAGAGTTT	341		
283	CAGAGCTGTCCGTATAAACTTCAGACAACTCTAAATCTGCAGCCCGAGCTAGGATAGAGCT	342		
342	CAGACGTGTCCGTATAAACTTCAGCAATCTTAAATCTGCAGCCCGAGCTAGGATAGAGCT	401		
343	TCATGAACCCCAATTTCAGAGGGAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCC	402		
402	TCATGAACCCCAATTTCAGAGGGAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCC	461		
403	AGACAGAGATGGAGACAAACTGCACCTTGCTCCACCCAGCCTGCCAAACAGTTTCTCAT	462		
462	AGACAGAGATGGAGACAAACTGCACCTTGCTCCACCCAGCCTGCCAAACAGTTTCTCAT	521		
463	CTCGCCCTTCTCCTCCCACTCTGGTGGCAGCCCATCAAGGATGCCACGCCAGTCTCT	522		
522	CTCGCCCTTCTCCTCCCACTCTGGTGGCAGCCCATCAAGGATGCCACGCCAGTCTCT	581		
523	CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCAAGGATGATGAGTCTCA	582		
582	CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCAAGGATGATGAGTCTCA	641		
583	TGCAGGACCTGAGTCCACCCCAAGTGTGCTGTGCAGCTGTGGCAGCTGCACATAGAGGA	642		
642	TGCAGGACCTGAGTCCACCCCAAGTGTGCTGTGCAGCTGTGGCAGCTGCACATAGAGGA	701		
643	AGAAGAGACCCCAAGACTTCCCAAGCCCAAAATCATCCAACTCGCGGTCTCGGCT	702		
702	AGAAGAGACCCCAAGACTTCCCAAGCCCAAAATCATCCAACTCGCGGTCTCGGCT	761		
703	GCCACCTTCGCTGTCCAACTGAGTGTGCTCTCTCGATATAGCCGTCTCTCTTT	762		
762	GCCACCTTCGCTGTCCAACTGAGTGTGCTCTCTCGATATAGCCGTCTCTCTTT	821		
763	ATCATGCTTTTTCCTCCCTGTGTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTTT	822		
822	ATCATGCTTTTTCCTCCCTGTGTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTTT	881		

[illegible]

RESULT 2
US-09-782-953-17
; Sequence 17, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In ver. 2.1

Query Match	44.2%	Score 413.2	DB 10	Length 594
Best Local Similarity	83.5%	Pred. No. 1.7e-125		
Matches 469	Conservative	0	Mismatches 93	Indels 0
QY	163	CTCGTTGTTGCGTGCATGTTCCACGAGTCAGTGTGTTGAAGGAGAGAGACGACGAGAAA	222	
Db	33	CACTCGTGCGCTGTGTGTGATGCGAGGTCTTTACCAATCAGGAGTTTAAAGAAA	92	
QY	223	ATTTAGGGAGCTGTTTCGGACTTATCATGACTGTGTGAGCTTCCAGCTATTTTAAAGAGTTT	282	
Db	93	ATTCGAGGGACTGTTCCGACCTATGATGAATGTGTGAGCTTCCAGCTGTTTAAAGAGTTT	152	
QY	283	CAGACGTGCCGTATATAACTTCAGCAATCCTAAATCTGCAGCCCGAGCTAGGATAGAGCT	342	
Db	153	CCGACGGGTTCGAATAAATTTTCAGGCATCCCAATCTGCAGCCGTGCCCGATAGAGCT	212	
QY	343	TCATGAAACCCAAATTCAGAGGGAAAAATTTAAGCTCTACTTTGCACAGGTTTCAGACTCC	402	
Db	213	TCATGAGACTCAGTTCAGAGGGAAGACTAAACTTACTTCGCCCAAGGTCAGACCCC	272	
QY	403	AGAGACATGGAGACAAACTGCATTTGGCTCCACCCAGGCTGCCAAACAGTTTCTCAT	462	
Db	273	AGACACATGGAGACAAACTGCATTTGGCACTCCACAGCTGCCAAACAGTTTCTCAT	332	
QY	463	CTCGGCCCTTCCTCCCACTGTGTGGCTGCAGCCCATCAACGATGCCAGCTCCT	522	
Db	333	CTCACCCCTTCATCTCTCTGTGTGGCTGGAAGCTATCAGCGATGCCACACAGCTCT	392	
QY	523	CAACTATGACTCTCTATGCTGTGGCCAAACTAGGACACGAGGAGAGAAATATGAGCTGCA	582	
Db	393	CAACTATGACTCTTATGCTGTGGCCAAACTAGGACACGAGGAGAGAAATATGAGCTGCA	452	
QY	583	TGACGGAGCTAGTCCACCCCAAGTGTGCTGTGGTGCAGTGTGGGACAGTCACATAGAGGA	642	
Db	453	CGTGGAACTAGTCTACACCGAGCGTGTGGTGCATGTGTGACAGGACATGGAGGA	512	
QY	643	AGAAGAGGACCCAAAGACTTCCCAAGCCAAAATATCATCCAACTCGGCTCCTGGCCT	702	
Db	513	GGAGGAGGACCCAAAGACTTCCCAAGCCAAAATATCATCCAACTCGGCTCCTGGCCT	572	
QY	703	GCCACCTCGGTGTCACATGA	724	
Db	573	GCCACCTCGGTGTCACATGA	594	

Query Match	29.9%	Score 279.2	DB 10	Length 828	
Best Local Similarity	66.0%	Pred. No. 3.5e-81			
Matches 420	Conservative	0	Mismatches 213	Indels 3	Gaps 1
QY	85	AGACAGGGGACTGGCGTGTACCTGGTTGTTTTGCAGAGAACGCCCTTTCAAGCAATCACTGA	144		
Db	85	AGAGAGGAAGAGAGATGATTTTGGTGAATAAGCATGATTGGATGAGATGATGGA	144		
QY	145	CTTCAATGACCTCCCAACTCGTTGTTTTGCGTGCAACTTCCACAGTCAGTGTGTTGAAGG	204		
Db	145	TTTAAGTGATCTGCGCTACTCTACITTTTGCTTGCAGCGCTCAGACAGTGTGTTGAGGC	204		
QY	205	AGAAGAGAGCAGGAGAAATTTGAGGACCTGTTCCGACTTATGATGACTGTGTGACGTT	264		
Db	205	ACGACAGCAGAGGAAGATTTGAAGCACTTCCACCATCTATGATGACCAGGTTACTTT	264		
QY	265	CCAGCTATTTAAGAGTTTCAGACGTGTCGTTATAAACTTCAGCAATCCTAAATCTCGAGC	324		
Db	265	TCAGCTGTTTAAAGCTTTAGACAGTCAAGATAAATTCAGCAAACTCTGAAGCGGCAGC	324		
QY	325	CCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGAAAAAATAAGCTCTACITT	384		
Db	325	AAGAGCGGAATAGAACTCCAGCAACAGACTTCAATGGCAGAACTAAAGCTATATTT	384		
QY	385	TGCACAGGTTTCAGACTCCAGACAGATGAGACAAACTGCACCTGGCTGCACCCAGCC	444		
Db	385	TGCACAGGTGCAGATGTCGGGGAAGTGGCGGACAGTCTATCTCTGCCGCCCGCAGCC	444		
QY	445	TGCCAAACAGTTTCATCTCGGCCCTTCTCTCCCACTGTGGCTGGCAGCCCATCAA	504		
Db	445	TGTCACAGCACTTCTCATCTCCCTCCAGCTCTCCCCAGTGGGTGGAAGCAGAGCGA	504		
QY	505	CGATGCCAGCCAGTCCCTCAACTATGACCTCTCTATGCTGGCGCAACTAGGACCAAG	564		
Db	505	AGATGGGATGCCCTGTTATAAATTATGATTTACTCTGTCTGTTTCCAAATTTGGGACCAAG	564		
QY	565	AGAGAAGTATGAGCTCCATGCAAGGACATGATGCCACCCCAAAGTGTGCTGTCAGCTGTG	624		
Db	565	AGAGAAATATGAACTTCAGCGCGGGAACAGAGTCGACACCCAGCGTGGTGTCTGTCG	624		
QY	625	CGACAGTGACATAGAGGAAGAGGAGCCCAAGACTTCCCCCAAGCCAAATAATCATCCA	684		
Db	625	TGAAGTGAACATGAAGAGGAAGAAGACAAAA--ACCCCAACAGAAAAATTGCCCA	681		
QY	685	AACTCGGCGTCTCGCCCTGCCACCTTCGGTGTCCAA	720		
Db	682	GACAGGCGCCCGACCTTCGACCGGAGCGTTGAA	717		

RESULT 5
US-09-782-953-23
; Sequence 23, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF
; TITLE OF INVENTION: CALCINEURIN INHIBITORS

us-10-030-613-3.rnpb

Tue Aug 26 09:11:05 2003

FILE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

FILE REFERENCE: US/09/782,953

CURRENT APPLICATION NUMBER: US/09/782,953

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 27

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 597

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(594)

US-09-782-953-2

Query Match 22.6%; Score 211; DB 10; Length 597;

Best Local Similarity 65.2%; Pred. No. 9,4e-59;

Matches 332; Conservative 0; Mismatches 165; Indels 12; Gaps 1;

143 GACTTCAATGACCTCCCAACTCGTTGTTGGTGGCAAGGCTTTCAGCAATCACTGA 144
 13 GATCTGAGGAGCTGCCGAGCGCCACCATCGCTCCGACCCGCGGTGTCGTG 72
 203 GGAGAAGAGAGCAAGGAAATTTGAGGAGCTGTTCCGGACTTATGATGACTGTGTGACG 262
 73 GACGCGCTGTGCCGGGCAATTTGAATCCCTCTTCAGAACATATGACAAGGACACCCACC 132
 263 TTCCAGCTATTTAAGAGTTTCAGAGCTGTCGGTATTAACCTTCAGCAATCTTAATCTTAC 322
 133 TTCCAGTATTTAAGAGCTTCAAACTGTCGGGATAAATTCAGCAACCCCTTATCTGCA 192
 323 GCCGAGCTAGATAGAGCTTTCATGAAACCAATTCAGAGGGAATAAATTAAGCTCTAC 382
 193 GCCGATGCCAGCTCGGCTGCACAGACCGAGTCTCTCCGCTTGGTGGGAAAGAAATCAAGTTGTAT 252
 383 TTTCACAGCTTCAGACTCCAGAGACAGATGAGAGACAACCTGCACCTTGGCTCCACCCAG 442
 253 TTTCCTCAGACTT-----TTACACATAGGAAGTTACACCTGGCTCCGCCCAAT 300
 443 CCTGCCAAGAGCTTTCATCTCGCCCTTCTCCACCTCTGTTGGCTGGCAGCCCATC 502
 301 CCGCAGAAACAGTTCTCATCTCCCTTCCCTCGGCTCTCTCCGCTTGGTGGGAAACAGTA 360
 503 AAGATGCCAGCCGCTCCTCAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCA 562
 361 GAAGATGCCAGCCGCTCATAAATACGATCTTTATATGCCATCTCCAGCTGGGGCCA 420
 563 GGAGAAGATGATGAGCTCCATGAGGAGCTGAGTCCACCCCAAGTGTCTGCTGACGTG 622
 421 GGAGAAGATGATGAGCTGCAATGAGGAGCAAGACCCCACTCCAGTGTGGTGTCCACGTG 480

RESULT 7

US-09-782-953-5

Sequence 5, Application US/09782953

Patent No. US20020150953A1

GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY

TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

FILE REFERENCE: US/09/782,953

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 27

FILE REFERENCE: US/09/782,953

CURRENT APPLICATION NUMBER: US/09/782,953

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 27

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 23

LENGTH: 720

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (2)..(637)

US-09-782-953-23

Query Match 24.8%; Score 231.6; DB 10; Length 720;

Best Local Similarity 63.5%; Pred. No. 1.7e-65;

Matches 404; Conservative 0; Mismatches 199; Indels 33; Gaps 2;

85 AGACAGGAGCTGGCTGTCACCTCGTTGTTGGCAAGGCTTTCAGCAATCACTGA 144
 7 AGAAGAGAGAGAGATGATTTTGGTGAATGAAGATGATTTGGATGATGATGA 66
 145 CTTCAATGACCTCCCAACTCGTTGTTGGCTGCAATGTTCCACAGTCACTGTTGAAGG 204
 67 TTTAAGTATCTGCTTACCTACCTTTTGGCTTGCAGCGTCCATGAAGCAGTGTGAGGC 126
 205 AGAAGAGAGCAAGGAAATTTGAGGAGCTGTTCCGACTTATGATGACTGTGTGACGTT 264
 127 ACGAGAGAGCAAGGAAATTTGAAGCACTCTTCACCACTATGATGACCACTTACTTT 186
 265 CCAGCTATTTAAGAGTTTCAGAGCTGTCGGTATTAACCTTCAGCAATCTTAATCTGCACG 324
 187 TCAGCTTTAAAGCTTTAAGAGAGTTCAGAAATTTTCAGCAATCTTAATCTGCACG 246
 325 CCGAGCTAGATGAGCTTTCATGAAACCAATTCAGAGGGAATAAATTAAGCTCTACTTT 384
 247 AAGAGCGGAGTATGAGCTTTCAGCAAGAGCTTCAATGCGCAGAACTAAGCTATATTT 306
 385 TGCACAGCTTCAGACTCCAGAGAGATGAGAGACAACCTGCATTTGGCTCCACCCAGCC 444
 307 TGCACAG-----TCTATCTCTCGCCGCCCCAGCC 336
 445 TGCACACAGTTTCTCATCTCGCCCTTCTCTCCACCTGTTGGCTGGCAGCCCATCAA 504
 337 TGTCAAGCAGTTCCTCATCTCCCTCCAGCTCTCCGCTGAGGAGGAGAGAGCGA 396
 505 CGATCCAGCGCAGTCTCAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCAAG 564
 397 AGATCGGATGCTGTTAATAATATGATTTACTCTGTGCTTTTCCAAATTTGGGACCAAG 456
 565 AGAGAAGTATGAGCTTCATGAGGAGCTGAGTCCACCCCAAGTGTCTGCTGCACTGTG 624
 457 AGAGAATATGAGCTTCACGCGGGAACAGAGTCGACACCCAGCTGGTGTGCTGCTG 516
 625 CGACAGTACATAGAGGAGAGAGAGGACCCCAAGAGCTTCCCAAGAGCCCAAAATCATCCA 684
 517 TGAAGTGAAGTGAAG 573
 685 AACTCGGCTGCTGGCTCCACCTCCCTCGTGTCCAA 720
 574 GACAAGGCGCCGCGCCCTCCGACCCGACGCGTGA 609

RESULT 6

US-09-782-953-2

Sequence 2, Application US/09782953

Patent No. US20020150953A1

GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

Tue Aug 26 09:11:05 2003

217 AAAGTGAACACGAGGCGCAAAATTTGAGTCCCTCTTTAGGACGCTATGACAAAGACATCACT 276
 264 TCCAGCTATTTAAGAGTTTTCAGAGTGTCCCTATATAAATTCAGCAATCCTTAATCTGCAG 323
 277 TTCAGTATTTTAAAGAGTTTCAAGAGTTCAGAAATTAATTCAGCAATCCTTAATCTGCAG 336
 324 CCGAGCTAGAGTATGAGCTTTCATGAACCCCAATTTAGAGGGGAAAAATTTAAAGCTCTACT 383
 337 CAGATGCCAGGCTCCAGCTGCATAGAGTGTGAGTTTCTGGGAAAGGAAATGAAGTTATAT 396
 384 TTGCACAGTTTCAGACTCCAGAGAGAGATGGAGCAAAATGCACATTTGGCTCCACCCAGC 443
 397 TTGCTCAG-----ACCTTACATATAGGAAGCTCAGCCTGCTCCGCGCAATC 444
 444 CTGCGCAACAGTTTCTCATCTCGCCCTCTCTCCACCTCTTGGCTGGCGACCCATCA 503
 445 CAGACAAAGAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAGTGG 504
 504 AGATGCCAGCCAGTCCATGACCTATGACCTCTCTATGCTGTGGCCAAACTAGGACCGAG 563
 505 AGATGCCAGCCAGTCCATGACCTATGACCTCTCTATGCTGTGGCCAAACTAGGACCGAG 564
 564 GAGAGAGTATGAGTCCATGACGAGGAGTCCACCCCAAGTGTGCTGCTGCAGCTGT 623
 565 GGGAAAGTATGATGATGACGAGGAGTCCACCCCAAGTGTGCTGCTGCAGCTGT 624
 624 GCGACAGTGCATAGAG-----CAAGAGAGGAGCCCAAAAGTCTCCCAAGC 671
 625 GTGAGAGTGCATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684
 672 CAAAGATCATCAAACTCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
 685 CAAATATATCCAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 743

RESULT 11

US-10-044-090-255

; Sequence 255, Application US/10044090

; Publication No. US20020137081A1

; GENERAL INFORMATION:

; APPLICANT: Olga Bandman

; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

; FILE REFERENCE: PA-0028 US

; CURRENT APPLICATION NUMBER: US/10/044,090

; CURRENT FILING DATE: 2002-01-09

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: PERL Program

; SEQ ID NO 255

; LENGTH: 2355

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1

; US-10-044-090-255

Query Match 20.0%; Score 187; DB 13; Length 2355;

Best Local Similarity 60.1%; Pred. No. 1.8e-50;

Matches 360; Conservative 0; Mismatches 215; Indels 24; Gaps 2;

144 ACTTCAATGACCTCCCACTCGTGTGTTGCGGCAATGTTCCAGCTATGATGACTGTGTGAGT 203
 164 ACTTCAATGACCTCCCACTCGTGTGTTGCGGCAATGTTCCAGCTATGATGACTGTGTGAGT 223
 204 GAGAGAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 263
 224 AAAGTGAACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 283
 264 TCCAGCTATTTAAGAGTTTTCAGAGTGTCCCTATATAAATTCAGCAATCCTTAATCTGCAG 323
 284 TTCAAGTATTTAAGAGTTTTCAGAGTGTCCCTATATAAATTCAGCAATCCTTAATCTGCAG 343
 324 CCCAGCTAGGATGAGAGTTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 383

203 GGAGAGAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 262
 97 GACGCGCTGTCCGCGGCAAAATTTGAGTCCCTCTTTAGGACGCTATGACAAAGACATCACT 156
 263 TTCAGCTATTTAAGAGTTTTCAGAGTGTCCCTATATAAATTCAGCAATCCTTAATCTGCAG 322
 157 TTCAGTATTTTAAAGAGTTTCAAGAGTTCAGAAATTAATTCAGCAATCCTTAATCTGCAG 216
 323 CCGGAGCTAGGATGAGCTTCATGAACCCCAATTTAGAGGGGAAAAATTTAAAGCTCTACT 382
 217 GAGATGCCAGGCTCCAGCTGCATAGAGTGTGAGTTTCTGGGAAAGGAAATGAAGTTATAT 276
 383 TTGACAGGTTTCAGACTCCAGAGAGAGATGGAGCAAAATGCACATTTGGCTCCACCCAGC 442
 277 TTGCTCAG-----ACCTTACATATAGGAAGCTCAGCCTGCTCCGCGCAAT 324
 443 CTGCGCAACAGTTTCTCATCTCGCCCTCTCTCCACCTCTTGGCTGGCGACCCATC 502
 325 CAGACAAAGAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAGTGG 384
 503 AACGATGCCAGGAGTCCATGACCTATGACCTCTCTATGCTGTGGCCAAACTAGGACCGA 562
 385 GAAGATGCCAGGAGTCCATGACCTATGACCTCTCTATGCTGTGGCCAAACTAGGACCGA 444
 563 GGAGAGAGTATGAGCTCCATGACGAGGAGTCCACCCCAAGTGTGCTGCTGCAGCTGT 622
 445 GGGAAAGTATGATGATGACGAGGAGTCCACCCCAAGTGTGCTGCTGCAGCTGT 504
 623 TGGCAGAGTGCATAGAG-----GAAGAGAGGAGCCCAAAAGTCTCCCAAGC 670
 505 TGTGAGAGTGCATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564
 671 CCAAAATATCAAACTCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
 565 CCAAAATATCCAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 624

RESULT 10

US-09-782-953-11

; Sequence 11, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; FILE REFERENCE: US/09/782,953

; CURRENT APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 11

; LENGTH: 2331

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (144)...(734)

; US-09-782-953-11

Query Match 20.0%; Score 187; DB 10; Length 2331;
 Best Local Similarity 60.1%; Pred. No. 1.8e-50;
 Matches 360; Conservative 0; Mismatches 215; Indels 24; Gaps 2;
 144 ACTTCAATGACCTCCCACTCGTGTGTTGCGGCAATGTTCCAGCTATGATGACTGTGTGAGT 203
 157 ACTTCAATGACCTCCCACTCGTGTGTTGCGGCAATGTTCCAGCTATGATGACTGTGTGAGT 216
 204 GAGAGAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 263

Tue Aug 26 09:11:05 2003

```

344 CAGTCATAAACTATGATCTCTTATATGCCATCTCTCAAGCTGGGGCCAGGGGAAAGATG 403
576 AGTCCATGAGGAGGAGTGTCCACCCCAAGTGTCTGTCAGCTGTGCGACAGTGTGACA 635
404 AATTGACGAGGAGGAGTGTCCACCCCAAGTGTCTGTCAGCTGTGCGACAGTGTGACA 463
636 TAGAG-----GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
464 AAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 523
684 AAATCGGGCTCTGCGCTGCGACCCCTCGTGTCCAACTGAGCTGCGC 730
524 AGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 570

```

```

RESULT 15
US-10-102-524-1746
; Sequence 1746, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Mannion, Paul A.
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1746
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-1746

```

```

Query Match      19.0%; Score 177.4; DB 14; Length 2173;
Best Local Similarity 62.0%; Pred. No. 2.5e-47;
Matches 327; Conservative 0; Mismatches 176; Indels 24; Gaps 2;

```

```

; QY      216 AGGAAAAATTTGAGGAGTCTTTGCGGACTTATGATGACTGTGTGACCTTCCAGCTATTTA 275
; DB      56 ATGCCAAATTTGAGTCTCTTTAGGACGATGATGACAGGACATCACCTTTTCAGTATTTA 115
; QY      276 AGAGTTTCAGAGTGTCTGCTATTAATCTGACAACTCTAAATCTGCGAGCCGAGCTAGGA 335
; DB      116 AGAGCTTCAACGAGTGTGAGTATTAATCTGACAACTCTTCTCCGAGCAGATGCCAGGC 175
; QY      336 TAGAGCTTCAATGAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 395
; DB      176 TCCAGCTGCATAGAGCTGAGTTTCTGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 231
; QY      396 AGACTCCAGAGAGAGTGGAGACAACTGGCTGCGCTGCGAGCCCACTGCGCAACAGT 455
; DB      232 -----ACCTTACATAGGAGGAGTCTGCGGAGTGGGATGGAACAGTGGGAGGAGGAGGAG 283
; QY      456 TTCTCATCTCCGCCCTTCTCTCCACCTGTTGGCTGGCAGCCCACTGCGAGTGGCAGGC 515
; DB      284 TTCTGATCTCCCTTCCCTTCTTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 343
; QY      516 CAGTCTCAACTATGAGCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
; DB      344 CAGTCTCAACTATGAGCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
; QY      576 AGCTCCATGCGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635
; DB      404 AATTGACGAGGAGGAGTGTCCACCCCAAGTGTCTGTCAGCTGTGCGACAGTGTGACA 463
; QY      636 TAGAG-----GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
; DB      464 AAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 523

```

```

413 CCAGCAAGAGGAGTGTCTCTCCCTCCGCTCTCCGSCAGTGGGATGGAACAAGT 472
503 AAGCATGCCACGCGAGTCTCAACTATGACTCTCTTATGCTGTGGCCAACTAGGACCA 562
473 GAAGATCGGAGCCGAGTCAATACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCA 532
563 GGAGAGAACTATGAGCTCCATGACGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
533 GGGGAAAGATGAAATGTCAGAGGAGTGTGACACCACTCCACGCTGGTGGTCCATGTA 592
623 TGGCAGAGTGCATAGAG-----GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 670
593 TGTAGAGTGTATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 652
671 CCAAAATATCAAACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
653 CCAAAATATCAAACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712

```

```

RESULT 14
US-09-880-107-3340
; Sequence 3340, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3340
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833
US-09-880-107-3340

```

```

Query Match      19.0%; Score 177.4; DB 10; Length 2173;
Best Local Similarity 62.0%; Pred. No. 2.5e-47;
Matches 327; Conservative 0; Mismatches 176; Indels 24; Gaps 2;

```

```

; QY      216 AGGAAAAATTTGAGGAGTCTTTGCGGACTTATGATGACTGTGTGACCTTCCAGCTATTTA 275
; DB      56 ATGCCAAATTTGAGTCTCTTTAGGACGATGATGACAGGACATCACCTTTTCAGTATTTA 115
; QY      276 AGAGTTTCAGAGTGTCTGCTATTAATCTGACAACTCTAAATCTGCGAGCCGAGCTAGGA 335
; DB      116 AGAGCTTCAACGAGTGTGAGTATTAATCTGACAACTCTTCTCCGAGCAGATGCCAGGC 175
; QY      336 TAGAGCTTCAATGAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 395
; DB      176 TCCAGCTGCATAGAGCTGAGTTTCTGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 231
; QY      396 AGACTCCAGAGAGAGTGGAGACAACTGCACTTGGCTTCCACCCCACTGCGCAACAGT 455
; DB      232 -----ACCTTACATAGGAGGAGTCTGCGGAGTGGGATGGAACAGTGGGAGGAGGAGGAG 283
; QY      456 TTCTCATCTCGGCCCTTCTCTCCACCTGTTGGCTGGCAGCCCACTGCGAGTGGCAGGC 515
; DB      284 TTCTGATCTCCCTTCCCTTCTTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 343
; QY      516 CAGTCTCAACTATGAGCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575

```

Page 9

us-10-030-613-3.rnpb

Qy
684 AAACCTGGCGTCTCCTGGCCCTGACCCTTCGGTGTCCAAGTGAGCTGCC 730

Dd
524 AGACAGAGGCGCGGAGTACAGCGCGATCCACCTCAGCTGAAGACTGGC 570

Search completed: August 24, 2003, 13:24:00
Job time : 326 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 09:53:05 ; Search time 2647 Seconds
(without alignments)
8575.890 Million cell updates/sec

Title: US-10-030-613-3
Perfect score: 934
Sequence: 1 agatcatattcatcgaa.....gtataaccattatttagaa 934

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_plp:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	705.2	75.5	753	10 AW957141	AW957141 EST369211
2	610.6	65.4	1201	9 AL532776	AL532776 AL532776
3	597.2	63.9	925	13 BQ720798	BQ720798 AGENCOURT
4	589.8	63.1	1055	12 BM925558	BM925558 AGENCOURT

5	581.6	62.3	1201	13 BX420406
6	531.8	56.9	557	9 AI796134
7	531.2	56.9	843	14 CA988986
8	516	55.2	516	9 AL693037
9	514.6	55.1	2094	11 AK049006
10	513.8	55.0	525	10 AW864635
11	494.6	53.0	661	14 CB531122
12	493.4	52.8	1007	10 BF534085
13	492.6	52.7	530	9 AA814235
14	480	51.4	749	13 BU847179
15	478	51.2	899	13 BU503652
16	474.4	50.8	873	13 BU285474
17	466.2	49.9	847	12 B1545609
18	460.8	49.3	538	14 CB457344
19	459.6	49.2	695	14 CB461803
20	440.4	47.2	596	14 CB538249
21	435.2	46.6	550	12 BG880881
22	421.8	45.2	485	9 AW655339
23	421.4	45.1	688	10 BG704949
24	421.4	45.1	729	13 BU313385
25	420.6	45.0	842	14 CB575637
26	414	44.3	758	10 BG709000
27	403.2	43.2	420	10 BF087797
28	403	43.1	474	9 AW470936
29	401	42.9	742	10 BG721994
30	395.6	42.4	676	10 BB635249
31	388.4	41.6	495	10 BF443461
32	388.4	41.6	748	12 BM964376
33	365.4	39.1	433	9 AA248590
34	359.4	38.5	660	9 AL892763
35	358.4	38.4	681	29 AG056729
36	356.8	38.2	726	29 AG059208
37	353.4	37.8	587	14 CB585503
38	349.4	37.4	626	10 BG701025
39	347.8	37.2	692	14 CB468653
40	346	37.0	536	14 CA894406
41	343.8	36.8	633	12 BM489970
42	336	36.0	338	14 T09144
43	334.4	35.8	767	13 BU611158
44	331.8	35.5	539	12 BG795691
45	329	35.2	347	12 BM994419

ALIGNMENTS

RESULT 1
AW957141
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AW957141
EST369211
AW957141
AW957141.1
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hedde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I.E., Saced, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
Unpublished
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnqu@tigr.org
Plate: 102
Seq primer: Reverse.
Location/Qualifiers

753 bp
mRNA
linear
EST 01-JUN-2000
MAGE resequences, MAGD Homo sapiens cDNA, mRNA sequence.

BX420406 BX420406
AI796134 WH42909.x
CA988986 AGENCOURT
AL693037 DKFZP313K
AK049006 Mus muscu
AW864635 PM4-SN001
CB531122 742372 MA
BF534085 602075073
AA814235 025H08.s
BU847179 AGENCOURT
BU503652 AGENCOURT
BU285474 603602128
B1545609 603187809
CB457344 714687 MA
CB461803 721815 MA
CB538249 776215 MA
BG880881 1b50f01.y
AW655339 h189d01.x
BG704949 602688927
BU313385 603540473
CB575637 AGENCOURT
BG709000 602675149
BF087797 IL3-HT061
AW470936 xz78c08.x
BG721994 602698760
BB635249 BB635249
BF443461 260951 MA
BM964376 UI-M-BQ0-
AA248590 csh0570.s
AL892763 AL892763
AG056729 Pan trogl
AG059208 Pan trogl
CB585503 AMGNNUC:N
BG701025 602682030
CB468653 734500 MA
CA894406 B0185C10-
BM489970 pqp2p.pk0
T09144 EST07037 In
BU611158 UI-M-FC0-
BG795691 UTSW_SM36
BM994419 UI-H-DH0-

[illegible]

Db 754 AAGAGAGGACCCAAAGACTTCCCAAGCCAAATAATCATCAAACTCGGGCTCTGGCC 813
 QY 702 TGCCACCTCGTGTCCAACTGAGCTGCTCTCTCTCGATAATAGCCGTCCTCTCTT 761
 Db 814 TGCCACCTCGTGTCCAACTGAGCTGCTCTCTCTCGATAATAGCCGTCCTCTCTT 873
 QY 762 TATCATGCTTTTCCCTCTGT-TTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGT 820
 Db 874 TATCATGCTTTTCCCTCTGTAGTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGT 933
 QY 821 TTGGTTGTTGAGATTCCTCTCTGTTTATCAAGCTCTCGGACAAAGAGGCTAGGAAAG 880
 Db 934 TTTGGTTGTTGAGATTCCTCTCTGTTTATCAAGCTCTCGGACAAAGAGGCTAGGAAAG 991
 QY 881 GTGATATGCTCTCTGATCATATATCAATCCAT 911
 Db 992 TGWATGTTCTGATMTWTATACATWAATAT 1022

RESULT 3

LOCUS B0720798 925 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT_8230839 Lupski_sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6191156 5', mRNA sequence.
 VERSION B0720798
 KEYWORDS EST.
 SOURCE B0720798.1 GI:21859695
 ORGANISM Homo sapiens (human)

REFERENCE

1 (bases 1 to 925)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1M13591 row: b column: 21
 High quality sequence start: 20
 High quality sequence stop: 702.

FEATURES

Source
 1..925
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6191156"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski_sympathetic_trunk"
 /note="vector: pCMV-Sport6 (Life Technologies); Site_1:
 Noti; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACGGCTCG-3' and
 5'-GACTAGTGTAGATCGGAGCGGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."
 BASE COUNT 227 a 266 c 211 g 221 t
 ORIGIN

Query Match

Best Local Similarity 63.9%; Score 597.2; DB 13; Length 925;
 92.1%; Pred. No. 4.5e-157;

Matches 652; Conservative 0; Mismatches 53; Indels 3; Gaps 2;
 QY 163 CTCGTTGTTTCCGTCGAATGTTCCACAGTCAGTGTGTAAGGAGAGAGCAAGAAAA 222
 Db 218 CACTCTGGTGTGCTGTGTGGATGTCGAGGTCTTTTACCAATCAGGAGGTTAAGGAAAA 277
 QY 223 ATTTGAGGAGCTGTTTTCGGACTTATGATGACGTGTGAGCTTCCAGCTATTTAAGAGTTT 282
 Db 278 ATTTGAGGAGCTGTTTTCGGACTTATGATGACGTGTGAGCTTCCAGCTATTTAAGAGTTT 337
 QY 283 CAGACGTGTCGATATAAATTCAGCAATCCTAAATCTGCAGCCGCGAGCTAGCATAGAGCT 342
 Db 338 CAGACGTGTCGATATAAATTCAGCAATCCTAAATCTGCAGCCGCGAGCTAGCATAGAGCT 397
 QY 343 TCATGAACCCCAATTCAGAGGAGGAAAAATTAAGCTCTTCTTTCACAGGTTTCAGAGTCC 402
 Db 398 TCATGAACCCCAATTCAGAGGAGGAAAAATTAAGCTCTTCTTTCACAGGTTTCAGAGTCC 457
 QY 403 AGAGACAGATGAGACAAACTGCTTCCACCCAGCCCTGCAACACAGTTTCTCAT 462
 Db 458 AGAGACAGATGAGACAAACTGCTTCCACCCAGCCCTGCAACACAGTTTCTCAT 517
 QY 463 CTCGCCCTTCTCTCCACCTGTTGGTGGCAGCCCATCAACAGATGCCGCCAGTCTCT 522
 Db 518 CTCGCCCTTCTCTCCACCTGTTGGTGGCAGCCCATCAACAGATGCCGCCAGTCTCT 577
 QY 523 CAACATGATGACCTCTCTATGCTGTGCCAAATAGGACGAGAGAGATGATGACTCCA 582
 Db 578 CAACATGATGACCTCTCTATGCTGTGCCAAATAGGACGAGAGAGATGATGACTCCA 637
 QY 593 TGCAGGAGCTGAGTCCACCCCAAGTGTCTGTGCACGTGTGCCACAGTGCATAGAGGA 642
 Db 638 TGCAGGAGCTGAGTCCACCCCAAGTGTCTGTGCACGTGTGCCACAGTGCATAGAGGA 697
 QY 643 AGAAGAGGACCAAGACTTCCCAAGCCCAAAATCATCAAACTGCGCTCTCTGGCT 702
 Db 698 AGAAGAGGACCAAGACTTCCCAAGCCCAAAATCATCAAACTGCGCTCTCTGGCT 757
 QY 703 GCCACCTCTGCTGCTCAACTGAGCTGCTCTCTCTCGATAATAGCCGTCCTCTCTT 762
 Db 758 GCCACCTCTGCTGCTCAACTGAGCTGCTCTCTCTCGAT-ATAGCCGCTCTCTCTT 816
 QY 763 ATCATGCTTTTCCCTCTGTTTGTTCARAAAAAATTCCTTTAAATTCCTGGGTGTTT 822
 Db 817 ATCATGCTTTTCCCTCTGTTTGTTCARAAAAAATTCCTTTAAATTCCTGGGTGTTT 876
 QY 823 GTTGTTTGAATTCCTTCC--TTGTTTATCAAGCTCTCGGACAAAAAG 868
 Db 877 GTTGTTTGAATTCCTTCCCTTGGTAATAAGGCTCTCGGACAAAAAG 924

RESULT 4
 LOCUS BM925558
 DEFINITION AGENCOURT_6710265 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5763757
 5', mRNA sequence.
 ACCESSION BM925558
 VERSION BM925558.1 GI:19375937
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1055)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

100

QY	239	CGAGCTTATGACTGTGTGACGTTCCAGCTATTAAAGAGTTTCAGAGCTGTCGGTATA	298
Db	570	: : : : : : : : : :	
QY	299	CGAGCTTATGACTGTGTGACGTTCCAGCTATTAAAGAGTTTCAGAGCTGTCGGTATA	629
QY	299	AAC TTCACAATCCTTAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTC	358
Db	630	: : : : : : : : : :	
QY	630	AAC TTCACAATCCTTAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTC	689
QY	359	AGAGGGAAAAAATTAAGACTCTACTTTGCACAGTTTCAGACTCCAGAGACAGATGGAGAC	418
Db	690	: : : : : : : : : :	
QY	690	AGAGGGAAAAAATTAAGACTCTACTTTGCACAGTTTCAGACTCCAGAGACAGATGGAGAC	749
QY	419	AAACTGCACCTTGCGTCCACC CGCTGCCAAACAGTTTTCATCTCGCCCCCTTCCTCC	478
Db	750	: : : : : : : : : :	
QY	750	AAACTGCACCTTGCGTCCACC CGCTGCCAAACAGTTTTCATCTCGCCCCCTTCCTCC	809
QY	479	CCACTGTGTGGCTGGCAGCCCATCAACGATGCCAGCGCAGTCTCACTATGACCTCCTC	538
Db	810	: : : : : : : : : :	
QY	810	CCACTGTGTGGCTGGCAGCSCTACAGATGCCACGCTCCTCAACTATGATCTCCTC	869
QY	539	TATGCTGTGGCCAACCTAGGACCAAGAGAAAGTATGAGCTCCATG CAGGAGCTGAGTCC	598
Db	870	: : : : : : : : : :	
QY	870	TATGCTGTGG - CAATAATAGCACGAGAGAAGTATGAGCTCCATG CAGGAGCTGAGTCC	928
QY	599	ACCCCAAGTGTGCTGTCGACAGTGTCCGACAGTGCACATAGAGGAAGAGGACCCAAAG	658
Db	929	: : : : : : : : : :	
QY	929	ACCCCAAGTGTGCTGTGCA - GTGTGGACAGTGACATAGAGGAA - AASAGGACCCCAAGA	986
QY	659	ACTTCCCACAAGCCAAAATCA	680
Db	987	: : : : : : : : : :	
Db	987	CTTCCC AAAAGCMAAAATCATCA	1008

RESULT 6

AI796134/c

LOCUS	DEFINITION
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

NOT INITIATED

ACCESSION

VERSION

KEYWORDS
SOURCE

SOURCE
ORGANISM

22

RESEARCH DESIGN

REFERENCE
AUTHORS

AUTHORS	TITLE
...	...

1

JOURNAL
COMPTON

RESULT 6
AI796134/C
LOCUS
AI796134
557 bp
linear
EST 06-JUL-1999
wh42g09.x1 NCI_CGAP_Kid1l Homo sapiens cDNA clone IMAGE:2383456 3',
similar to TR:Q14206 Q14206 ZAKI-4 PROTEIN. ;, mRNA sequence.
AI796134
ACCESSION

FEATURES

Query Match	56.9%;	Score 531.2;	DB 14;	Length 843;
Best Local Similarity	83.9%;	Pred. No. 1.8e-138;		
		Mismatches 120;		
		Indels 4;	Gaps 4;	

Matches	646;	Conservative	U; Mismatches	120; Indels
Qy	34	GCACCAGGACACTCCCTGAAGATGGAGGACTTTTCTTACTGTGTGTCATAGACAGGA	93	

Qy 154 C T C C C C C A A C T C G T T G T T T G C G T C A A T G T T T C A C C A G T C A G T G T T T G A A G G A G A G A G A G 213

	DB	264	TAAGAGTTTCCGACGGGTTCGAATAAATTCAGCCATCCCAAAATCTCAGGCCCGCCGCCCCG	393
	QV	334	GATAGAGCTTCATGAACCACCAATTTCAGAGGGAAAAAATTAAGACTCTACTTTGCACAGGT	

324 GATAGAGCTTCATGAGACTCAGTTCAGAGGGGAAGAGCTTAAACTCTACTTCGCCAGGT 383

Db

[illegible]

QY 454 GTTCTCATCTCGCCCCCTTCCCTCCACACCTGTGGCTGGCAGCCCATCAACGATGCCAC 513

	Db	444	GTTCCTCATCTCACCCCCCTTCACTCCTCTCTGTTGGCTGGAAAGCCCTATACAGCGAAGGCCAC	573
	Ov	514	GCCAGTCCCTCAACTATGACCTCCTCTATGCTGTGGCCAACTAGGACCAGGAGAGAGAAGTA	573

504 ACCAGTCTCAACTATGACCTCTTTATGCTGTGGCCAAACTAGGACCAAGAGAGAAATA 563

QY	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800
QY	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	79			


```

|||||
181 TTGGCTGGCAGCCCATCAACGATGCCAGCCAGTCCCTCAACTATGACCTCTCATGCTG 240
QY TGGCCAAACTAGGACAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 605
Db TGGCCAAACTAGGACAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 241
QY GTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 606
Db GTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 301
QY 666 CAAAGCAAAATATCACTCAAACTCGGCTGCTGCGCTGCCACCTCGCTGCCACCTCG 725
Db 361 CAAAGCAAAATATCACTCAAACTCGGCTGCTGCGCTGCCACCTCGCTGCCACCTCG 420
QY 726 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 785
Db 421 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 786 TTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTT 821
Db 481 TTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTT 516

```

```

RESULT 9
LOCUS AK049006
DEFINITION Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:C230090N07 product:Down syndrome critical
region gene 1-like 1, full insert sequence.
ACCESSION AK049006
VERSION AK049006.1 GI:26339679
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1

```

```

REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,H., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,H., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,I.,
Saito,T., Okazaki,Y., Gojohori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staabli,F., Suzuki,R., Tomita,M.,

```

Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mommaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2094)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/.

Location/Qualifiers

1. 2094

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:C230090N07"

/db_xref="taxon.10090"

/clone="C230090N07"

/tissue_type="cerebellum"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

200. 793

/note="unnamed protein product; Down syndrome critical

region gene 1-like 1 (MGDINGI:1858219, GBINA_030598,

evidence: BLASTN, 100%, match=659)

putative"

/codon_start=1

/protein_id="BAC33511.1"

/db_xref="GI:26339680"

/translation="MPAPSDCDVSTLVACVDVVEFTNQEVKEFGLRTYDECVT
FOLFKSRFRIRNFHPKSAARARIELHETQFGKKLYFAQVQVPTDGGKHLAP
POPAPQFLISPPSPVGVKPISDATPVLYNLLYAVAKLGPCKYELHAGTESTPV
VYHVCDSMEEEDPKTSPKPKIIQTRRGLPPSVSN"

CDS

FEATURES

SOURCE

361 AGGGAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGACAGATGGAGACAA 420

3 ATTTGAGGGAC TGTTTCGGACTTATGATGACTGTGTGACTTCCAGCTATTTAACAGTTT 282

150	ATTCCAGGAGCTGTTCGGGACCTATGATGAAATGTTGACGTTCCAGCTGTTTAAAGAGTTT	209
283	CAGACGCTGTCGTATAAACTTCAGCAATCCTAAATCTGAGCCCGAGCTAGGATAGAGCT	342
210	CGACGGGTTTCCAAATAATTTTCAGCCATCCCAAACTGACGCCGTGCCCGSATAGAGCT	269
343	TCATGAACCCAAATTCAGAGGGGAAAAAATAAGCTCTACTTTTGACACAGGTTTCAGACTCC	402
270	TCATGAGACTTCAGTTTCAGAGGGAAGAAGCTAAAACTCTACTTCGCCCGAGTCAGACCCC	329
403	AGAGACAGATGGAGACAAACTGACACTTGGCTCCACCCCGAGCTGCCAAAACAGTTTCTCAT	462
330	AGAGACAGATGGAGACAAACTGCAATTTGGCACCTCCACAGCTTGCCAAAACAGTTCTCAT	389
463	CTCGCCCCCTTCCTCCCGACCTGTTGGCTGGCAGGCCCATCAAGCATGCCACGCCAGTCCT	522
390	CTCACCCCTTTCATCTCCCTCTTGGCTGGGAAGCCTATTCAGCGATGCCACACCACTCT	449
523	CAACTATGACCTCCTCTATGCTGTGGCCAACTAGGACACGAGAGAAAGTATGAGCTCCA	582
450	CAACTATGACCTTCTTTTATGCTGTGGCCAACTAGGACACGAGAGAAATATGAGCTGCA	509
583	TGCAGGACTGAGTGCACCCCAAGTCGCTGTCACGCTGTGCCACAGTGCACATAGAGGA	642
510	CGCTGAACTGAGTCTACCCGAGCTGCTGTTGTCATGTGTGACAGCGACATGGAGGA	569
643	AGAAGAGGACCCAAAGACTTCCCCAAAGCCAAAATCATCCAACTCGGCGTCTCTGGCCT	702
570	GGAGGAGGACCCAAAGACTTCCCCAAAGCCAAAATCATCCAGACCCGGCGTCCCGGCT	629
703	GCCACCTCCGTGTCCAACTGAGG - TGCCTGCTCCTTCTCGATATAATAGCCGCTCTCTCTT	761
630	GCCACCTCCGTGTCCAACTGAGGACGCTCCATGATACATGCACTCTCTCTATCATGCT	689
762	TATCATGCTTTTTCCTCCCTGTGTTTGTCAAAAAAATTCGCTTTAAATTCCTGGGT - GT	820
690	TCCGCCCCCCCCCTTGTGTTTGTCCAAACAATTAATTTGCTTTAAGTCCCTGGGTGGT	749
821	TTGGTTGTTTGAGATTCTTCTCTGTTATCAAGCCTCTCGGACAAAAAGGCTAGGAAAAG	880
750	TTGGTTGGTTGCGAATCCCTCCTTGTAAATTANGCCTTCGGACAAAAGGCTAGGCAAG	809
881	GTGATATGCTCTCTGGA - TCATATCATACCCATTAAAGTATAACCCATTATTT	930
810	GTGATATGCTCTCCAAATCACTCCCTACCCATTAACTACAAATCCATCACTT	861

Search completed: August 24, 2003, 12:11:21
Job time : 2652 secs